

Simultaneous confidence bands for contrasts between several nonlinear regression curves

Xiaolei Lu^{*} and Satoshi Kuriki[†]

Abstract

We propose simultaneous confidence bands of the hyperbolic-type for the contrasts between several nonlinear (curvilinear) regression curves. The critical value of a confidence band is determined from the distribution of the maximum of a chi-square random process defined on the domain of explanatory variables. We use the volume-of-tube method to derive an upper tail probability formula of the maximum of a chi-square random process, which is sufficiently accurate in commonly used tail regions. Moreover, we prove that the formula obtained is equivalent to the expectation of the Euler-Poincaré characteristic of the excursion set of the chi-square random process, and hence conservative. This result is therefore a generalization of Naiman's inequality for Gaussian random processes. As an illustrative example, growth curves of consomic mice are analyzed.

Keywords and phrases. Chi-square random process, expected Euler-characteristic heuristic, Gaussian random field, growth curve, Naiman's inequality, volume-of-tube method.

1 Introduction

This paper concerns multiple comparisons of k (≥ 3) nonlinear (curvilinear) regression curves estimated from independent k groups. Suppose that for each group $i = 1, \dots, k$, and for each explanatory variable $x_j \in \mathcal{X}$, $j = 1, \dots, n$, we have observations $y_{ij1}, \dots, y_{ijr_i}$ as objective variables with r_i replications, which are assumed to follow the model

$$y_{ijh} = g_i(x_j) + \varepsilon_{ijh}, \quad i = 1, \dots, k, \quad j = 1, \dots, n, \quad h = 1, \dots, r_i. \quad (1.1)$$

Here, $\mathcal{X} \subseteq \mathbb{R}$ is the domain of explanatory variables, and random errors ε_{ijh} are assumed to be independently distributed as the normal distribution $N(0, \sigma(x_j)^2)$. The variance function $\sigma(x)^2$ is supposed to be known, or at least known up to a constant $\sigma(x)^2 =$

^{*}Department of Statistical Sciences, Graduate University for Advanced Studies, 10-3 Midoricho, Tachikawa, Tokyo 190-8562, Japan, Email: lu.xiaolei@ism.ac.jp

[†]The Institute of Statistical Mathematics, 10-3 Midoricho, Tachikawa, Tokyo 190-8562, Japan, Email: kuriki@ism.ac.jp

$\sigma^2\sigma_0(x)^2$. In the case of the latter, we suppose that an independent estimator $\hat{\sigma}^2$ of σ^2 is available. In addition, we assume that the true regression curve has the form

$$g_i(x) = \beta_i^T f(x), \quad x \in \mathcal{X}, \quad (1.2)$$

where $f(x) = (f_1(x), \dots, f_p(x))^T$ is a known regression basis vector function, $\beta_i = (\beta_{i1}, \dots, \beta_{ip})^T$ is an unknown parameter vector. Then, the least squares estimator $\hat{\beta}_i$ of β_i has the multivariate normal distribution $N_p(\beta_i, r_i^{-1}\Sigma)$, where

$$\Sigma = \left(\sum_{j=1}^n \frac{1}{\sigma(x_j)^2} f(x_j) f(x_j)^T \right)^{-1}$$

is the inverse of the $p \times p$ information matrix. When $\sigma(x)^2 = \sigma^2\sigma_0(x)^2$, we have $\Sigma = \sigma^2\Sigma_0$, where Σ_0 is Σ with $\sigma(x_j)$ replaced by $\sigma_0(x_j)$.

Let \mathcal{C} denote the set of vectors $c = (c_1, \dots, c_k)^T$ such that $\sum_{i=1}^k c_i = 0$. The focus of this paper is the construction of $1 - \alpha$ simultaneous confidence bands for all the contrasts $\sum_{i=1}^k c_i g_i(x) = \sum_{i=1}^k c_i \beta_i^T f(x)$ between the k regression curves for all $x \in \mathcal{X}$ and $c \in \mathcal{C}$, where \mathcal{X} is a given finite interval $[a, b]$, a finite union of intervals $\bigsqcup_i [a_i, b_i]$, or an infinite interval $(-\infty, \infty)$, with the symbol ' \bigsqcup ' denoting disjoint union. Specifically, according to the traditional form of the point estimate plus or minus a probability point times the estimated standard error, we construct a $1 - \alpha$ simultaneous confidence band of the form

$$\sum_{i=1}^k c_i \beta_i^T f(x) \in \sum_{i=1}^k c_i \hat{\beta}_i^T f(x) \pm b_{1-\alpha} \sqrt{\left(\sum_{i=1}^k \frac{c_i^2}{r_i} \right) f(x)^T \Sigma f(x)}, \quad (1.3)$$

where $\hat{\beta}_i^T f(x)$ is the estimator of $\beta_i^T f(x)$ in (1.2). This form is referred to as a hyperbolic-type (Liu, 2010). The critical value $b_{1-\alpha}$ is determined such that the event in (1.3) for all $x \in \mathcal{X}$ and $c \in \mathcal{C}$ holds with a probability of at least $1 - \alpha$. Our problem typically arises from growth curve analysis and longitudinal data analysis.

Throughout this paper, we assume that the regression curve $g_i(x)$ is a linear combination of a finite number of known basis functions in (1.2). Although it is a conventional regression model, we must always be careful about the approximation bias caused by model misspecification. This issue is examined in Section 5.

The problem concerning the construction of simultaneous confidence bands in a regression model originates with Working and Hotelling (1929). They formalized this problem as the construction of confidence intervals for an estimated regression line, and provided a critical value by making use of the Cauchy-Schwartz inequality. Specifically, Working and Hotelling (1929) treated the case of

- (i) one regression model (equivalent to the case $k = 2$ in our problem),
- (ii) the simple regression $f(x) = (1, x)^T$, and
- (iii) the unrestricted domain of the explanatory variables $\mathcal{X} = (-\infty, \infty)$.

Subsequently, many reports concerning the relaxation of these conditions have appeared in the literature.

In the case of one regression model, Wynn and Bloomfield (1971) pointed out that the use of the Cauchy-Schwartz inequality leads to conservative bands unless both (ii) and (iii) hold. They illustrated improved confidence bands for the quadratic regression $f(x) = (1, x, x^2)^T$. Uusipaikka (1983) constructed exact confidence bands for linear regressions when \mathcal{X} is a finite interval. See Liu, Lin, and Piegorsch (2008) and Liu (2010) for historical reviews. The problem of comparisons of $k \geq 3$ regression curves was considered by Spurrier (1999, 2002) and Lu and Chen (2009), who proposed procedures based on simple linear regressions. However, it is difficult to extend these methods to nonlinear regressions.

One exception is Naiman (1986)'s integral-geometric approach. In the unit sphere \mathbb{S}^{p-1} of the p -dimensional Euclidean space, he defined a trajectory

$$\Gamma = \overline{\{\psi(x) \mid x \in \mathcal{X}\}} \subset \mathbb{S}^{p-1} \quad (1.4)$$

of a normalized basis vector function

$$\psi(x) = \frac{\Sigma^{\frac{1}{2}} f(x)}{\|\Sigma^{\frac{1}{2}} f(x)\|}, \quad (1.5)$$

and evaluated the volume of the tubular neighborhood of Γ . In the case of one regression model, he constructed a simultaneous confidence band with the critical value obtained from this volume. The volume formula for such tubes originates from Hotelling (1939) and Weyl (1939). Currently, this idea is understood in the framework of the volume-of-tube method (Sun (1993), Kuriki and Takemura (2001, 2009), Takemura and Kuriki (2002), Adler and Taylor (2007)). As shown in Section 2, we require the tail probability of the maximum of a Gaussian random field or a chi-square random process as a pivotal quantity. The volume-of-tube method is a methodology to evaluate such tail probabilities.

In this paper, we adopt this integral-geometric approach. In the case of $k \geq 3$, we define a subset M in (3.1) of a unit sphere, and by evaluating the volume of its tubular neighborhood, we obtain the critical value $b_{1-\alpha}$ in (1.3) by means of the volume-of-tube method. Moreover, we prove that the proposed confidence band is conservative. It is known that Naiman (1986)'s confidence band is conservative (Naiman's inequality, see also Johnstone and Siegmund (1989)), and our result is regarded as its generalization.

Note that, in the setting of this paper, the covariance matrices of the estimators $\hat{\beta}_i$ are identical up to a multiplicative constant. This property arises from the condition that the explanatory variables x_j are common between k groups in the model (1.1). This represents the so-called the balanced case. For the unbalanced case, the problem of constructing simultaneous confidence bands is quite tedious and only simulation-based approaches are available (Liu, Jamshidian, and Zhang (2004), Liu, Wynn, and Hayter (2008), Jamshidian, Liu, and Br (2010), Liu (2010)). In this paper, we address only the balanced case.

Note also that, in the one-group case ($k = 1$), various simultaneous confidence bands have been proposed by means of the volume-of-tube method. Johansen and Johnstone (1990) demonstrated the usefulness of Hotelling's volume formula for the construction of

simultaneous bands. The application to the B-spline regression is found in Shen, Wolfe, and Zhou (1998). Sun and Loader (1994) proposed a modification to the volume-of-tube formula when a small approximation bias caused by model misspecification exists. In succeeding papers, Sun and her coauthors developed this idea in various model settings (Faraway and Sun (1995), Sun, Raz, and Faraway (1999), Sun, Loader, and McCormick (2000)). See also Krivobokova, Kneib, and Claeskens (2010). The crucial difference between this paper and existing work is that in this paper we need to treat a Gaussian random field with a general dimensional ($k - 1$ dimensional) index set, and need the volume formula up to an arbitrary order.

The layout of the paper is as follows. In Section 2 we define a Gaussian random field and a chi-square random process as pivotal quantities. We show that the critical value $b_{1-\alpha}$ is determined from the upper tail probability of the maximum of a Gaussian random field or a chi-square random process. In Section 3, the volume-of-tube method and its related method referred to as the expected Euler-characteristic heuristics are briefly summarized. The main results are provided in Section 4. Some simulation study under model misspecification is conducted in Section 5. Section 6 is devoted to the analysis of growth curve data. Details of the proofs are in the Appendix.

2 Random fields as pivotal quantities

Our problem is to determine the critical value $b_{1-\alpha}$ in (1.3). Suppose first that Σ is fully known. Define a pivotal quantity

$$T(x, c) = \frac{\sum_{i=1}^k c_i (\hat{\beta}_i - \beta_i)^T f(x)}{\sqrt{\left(\sum_{i=1}^k \frac{c_i^2}{r_i}\right) f(x)^T \Sigma f(x)}}. \quad (2.1)$$

Then, the critical value $b_{1-\alpha}$ is the solution b of the equation

$$P(T(x, c) \leq b, \forall x \in \mathcal{X}, \forall c \in \mathcal{C}) = P\left(\max_{x \in \mathcal{X}, c \in \mathcal{C}} T(x, c) \leq b\right) = 1 - \alpha.$$

In this expression, we use $T(x, c)$ instead of $|T(x, c)|$, since $c \in \mathcal{C}$ implies $-c \in \mathcal{C}$ and $|T(x, c)|$ is equal to $T(x, c)$ or $T(x, -c)$. Inverting $|T(c, x)| \leq b_{1-\alpha}$ yields the $1 - \alpha$ simultaneous confidence band in (1.3).

In the following, we show that $b_{1-\alpha}^2$ is the upper α point of the maximum of a chi-square random process. We can assume that $\sum_{i=1}^k c_i^2/r_i = 1$ without loss of generality, because $T(x, c)$ is a homogeneous function in c . Let $\rho = (\sqrt{r_1}, \dots, \sqrt{r_k})^T$, and define a $k \times (k - 1)$ matrix H such that $\rho^T H = 0$, $H^T H = I_{k-1}$, and $HH^T = I_k - \rho\rho^T/(\rho^T \rho)$. (An example of H is given in Remark 2.1 below.) Then the $c = (c_1, \dots, c_k)^T$ such that $\sum_{i=1}^k c_i^2/r_i = 1$ and $\sum_{i=1}^k c_i = 0$ are represented as

$$c = \text{diag}(\sqrt{r_1}, \dots, \sqrt{r_k}) H h, \quad h \in \mathbb{S}^{k-2},$$

where \mathbb{S}^{k-2} is the set of $(k-1)$ -dimensional unit column vectors.

Let $\Sigma^{\frac{1}{2}}$ be a matrix such that $(\Sigma^{\frac{1}{2}})^T \Sigma^{\frac{1}{2}} = \Sigma$, and let $\Sigma^{-\frac{1}{2}}$ be its inverse. Then, $\eta_i = \sqrt{r_i}(\Sigma^{-\frac{1}{2}})^T(\hat{\beta}_i - \beta_i)$ is distributed normally as $N_p(0, I)$, independently for $i = 1, \dots, k$. Let $\psi : \mathcal{X} \rightarrow \mathbb{S}^{p-1}$ as defined in (1.5). Then, $T(x, c)$ is rewritten as

$$\begin{aligned}
T(x, c) &= \sum_{i=1}^k \frac{c_i}{\sqrt{r_i}} \sqrt{r_i} ((\Sigma^{-\frac{1}{2}})^T (\hat{\beta}_i - \beta_i))^T \frac{\Sigma^{\frac{1}{2}} f(x)}{\|\Sigma^{\frac{1}{2}} f(x)\|} \\
&= c^T \text{diag}(\sqrt{r_1}, \dots, \sqrt{r_k})^{-1} \begin{pmatrix} \eta_1^T \\ \vdots \\ \eta_k^T \end{pmatrix}_{k \times p} \psi(x) \\
&= h^T \begin{pmatrix} \xi_1^T \\ \vdots \\ \xi_{k-1}^T \end{pmatrix}_{(k-1) \times p} \psi(x) \\
&= \xi^T (h \otimes \psi(x)), \tag{2.2}
\end{aligned}$$

where ξ_i are $p \times 1$ vectors defined by $(\xi_1, \dots, \xi_{k-1})_{p \times (k-1)} = (\eta_1, \dots, \eta_k)_{p \times k} H$, $\xi = (\xi_1^T, \dots, \xi_{k-1}^T)^T$ is a $p(k-1) \times 1$ vector, and ‘ \otimes ’ is the Kronecker product. Because the vectors η_i consist of independent standard Gaussian random variables $N(0, 1)$, so does the vector ξ . When x and h are fixed, because of $\|\psi(x)\| = \|h \otimes \psi(x)\| = 1$, $\xi_i^T \psi(x)$ is distributed as $N(0, 1)$ independently for $i = 1, \dots, k$, and $\xi^T (h \otimes \psi(x))$ is distributed as $N(0, 1)$.

From (2.2), we can see that

$$\max_{c \in \mathcal{C}} T(x, c) = \sqrt{\sum_{i=1}^{k-1} (\xi_i^T \psi(x))^2}. \tag{2.3}$$

For each fixed x , this is distributed as the square root of the chi-square distribution χ_{k-1}^2 with $k-1$ degrees of freedom.

When $\Sigma = \sigma^2 \Sigma_0$ with Σ_0 known, and an independent estimator $\hat{\sigma}^2 \sim \sigma^2 \chi_{\nu}^2 / \nu$ of unknown σ^2 is available, we redefine $T(x, c)$ in (2.1) by replacing Σ in the denominator with $\hat{\sigma}^2 \Sigma_0$, and instead of (2.2) and (2.3) we have

$$T(x, c) = \frac{1}{\tau} \xi^T (h \otimes \psi(x)), \quad \max_{c \in \mathcal{C}} T(x, c) = \sqrt{\frac{1}{\tau^2} \sum_{i=1}^{k-1} (\xi_i^T \psi(x))^2}, \quad \tau^2 = \frac{\hat{\sigma}^2}{\sigma^2}.$$

Now we consider the object in (2.2) as a random function of (x, h) :

$$Z(x, h) = \xi^T (h \otimes \psi(x)), \quad (x, h) \in \mathcal{X} \times \mathbb{S}^{k-2}, \tag{2.4}$$

where $\xi \sim N_{p(k-1)}(0, I)$. Then, $Z(x, h)$ is the Gaussian random field with mean 0, variance 1, and a covariance function

$$\text{Cov}(Z(x, h), Z(\tilde{x}, \tilde{h})) = \psi(x)^T \psi(\tilde{x}) \cdot h^T \tilde{h}.$$

Similarly, we define the chi-square random process with $k - 1$ degrees of freedom:

$$Y(x) = \sum_{i=1}^{k-1} (\xi_i^T \psi(x))^2, \quad x \in \mathcal{X}. \quad (2.5)$$

We summarize the results of this section below.

Theorem 2.1. *When Σ is known, the critical value $b_{1-\alpha}$ is determined as the solution $b = b_{1-\alpha}$ of*

$$P\left(\max_{x \in \mathcal{X}, h \in \mathbb{S}^{k-2}} Z(x, h) \geq b\right) = P\left(\max_{x \in \mathcal{X}} Y(x) \geq b^2\right) = \alpha,$$

where $Z(x, h)$ is the Gaussian random field defined in (2.4), and $Y(x)$ is the chi-square random process defined in (2.5).

When $\Sigma = \sigma^2 \Sigma_0$ with Σ_0 known, the critical value $b_{1-\alpha}$ is determined as the solution $b = b_{1-\alpha}$ of

$$E\left[P\left(\max_{x \in \mathcal{X}, h \in \mathbb{S}^{k-2}} Z(x, h) \geq b\tau \mid \tau^2\right)\right] = E\left[P\left(\max_{x \in \mathcal{X}} Y(x) \geq b^2\tau^2 \mid \tau^2\right)\right] = \alpha,$$

where the expectation is taken over $\tau^2 \sim \chi_\nu^2/\nu$, with ν being the degrees of freedom of the estimator of σ^2 .

Remark 2.1. *An example of $k \times (k - 1)$ matrix H such that $\rho^T H = 0$, $H^T H = I_{k-1}$, $H H^T = I_k - \rho \rho^T / (\rho^T \rho)$ with $\rho = (\sqrt{r_1}, \dots, \sqrt{r_k})^T$ is given as*

$$H = \begin{pmatrix} \frac{\sqrt{r_1 r_2}}{\sqrt{R_1 R_2}} & \frac{\sqrt{r_1 r_3}}{\sqrt{R_2 R_3}} & \cdots & \frac{\sqrt{r_1 r_k}}{\sqrt{R_{k-1} R_k}} \\ -\frac{R_1}{\sqrt{R_1 R_2}} & \frac{\sqrt{r_2 r_3}}{\sqrt{R_2 R_3}} & \cdots & \frac{\sqrt{r_2 r_k}}{\sqrt{R_{k-1} R_k}} \\ & -\frac{R_2}{\sqrt{R_2 R_3}} & \cdots & \frac{\sqrt{r_3 r_k}}{\sqrt{R_{k-1} R_k}} \\ & & \ddots & \vdots \\ 0 & & & -\frac{R_{k-1}}{\sqrt{R_{k-1} R_k}} \end{pmatrix}_{k \times (k-1)},$$

where $R_i = \sum_{j=1}^i r_j$.

3 Preliminaries on the volume-of-tube method

In this section, we summarize the volume-of-tube method for evaluating the upper tail probability of the maximum of a Gaussian random field.

Let ξ be a Gaussian random vector distributed as $N_n(0, I)$. Let M be a closed subset of \mathbb{S}^{n-1} , the unit sphere (the set of unit column vectors) of \mathbb{R}^n . Then, the random map $u \mapsto \xi^T u$, $u \in M$, is a Gaussian random field with mean 0, variance 1, and a covariance function $\text{Cov}(\xi^T u, \xi^T v) = u^T v$. The volume-of-tube method approximates the distribution

of the maximum $\max_{u \in M} \xi^T u$. The maximum of $Z(x, h)$ in (2.4) can be treated in this framework by setting

$$M = \overline{\{h \otimes \psi(x) \mid (x, h) \in \mathcal{X} \times \mathbb{S}^{k-2}\}} \quad \text{and} \quad n = p(k-1). \quad (3.1)$$

The dimension of M is $d = \dim M = k-1$. To apply the volume-of-tube method, we require the following assumption on M .

Assumption 3.1. *M is a d -dimensional closed piecewise C^2 -manifold, or M is a d -dimensional C^2 -manifold with piecewise C^2 -boundary. We write $M = \text{Int}M \sqcup \partial M$, where $\text{Int}M$ and ∂M denote the interior and the boundary of M , respectively. In the former case, $\partial M = \emptyset$.*

When M is defined by (3.1), we can provide a sufficient condition for Assumption 3.1.

Assumption 3.2. *$\psi : \mathcal{X} \rightarrow \mathbb{S}^{p-1}$ is a one-to-one map of class piecewise C^2 . There does not exist $x, \tilde{x} \in \mathcal{X}$ such that $\psi(x) = -\psi(\tilde{x})$.*

Under Assumption 3.2, the map $(x, h) \mapsto h \otimes \psi(x)$ is a piecewise C^2 one-to-one map.

Example 3.1. *Consider the polynomial regression with a basis function vector $f(x) = (1, x, \dots, x^{p-1})^T$. When the domain of x is a finite interval $\mathcal{X} = [a, b]$, we have*

$$\begin{aligned} \text{Int}M &= \{h \otimes \psi(x) \mid x \in (a, b), h \in \mathbb{S}^{n-1}\}, \\ \partial M &= \{h \otimes \psi(a) \mid h \in \mathbb{S}^{n-1}\} \sqcup \{h \otimes \psi(b) \mid h \in \mathbb{S}^{n-1}\}. \end{aligned}$$

When $\mathcal{X} = (-\infty, \infty)$, $\psi(\pm\infty) = \pm \Sigma_p^{\frac{1}{2}} e_p / \sqrt{e_p^T \Sigma e_p}$ with $e_p = (0, \dots, 0, 1)^T$, and hence $h \otimes \psi(\infty) = ((-1)^{p-1} h) \otimes \psi(-\infty)$. This means that M is a closed manifold without boundary.

Example 3.2. *Consider the trigonometric regression with a basis function vector*

$$f(x) = \left(1, \sqrt{2} \cos x, \sqrt{2} \sin x, \dots, \sqrt{2} \cos mx, \sqrt{2} \sin mx\right)^T.$$

When $\mathcal{X} = [0, 2\pi)$, M is a closed manifold without boundary.

We now define “tube”, the key notion of the volume-of-tube method. The set of points of \mathbb{S}^{n-1} whose great circle distance from M is less than or equal to θ is the tube about M with radius θ , and has the expression

$$M_\theta = \left\{v \in \mathbb{S}^{n-1} \mid \min_{u \in M} \cos^{-1}(u^T v) \leq \theta\right\}.$$

If the radius θ is sufficiently small, the tube M_θ does not have self-overlap, whereas when the radius θ is large, the tube has self-overlap. The threshold radius between the two cases is known as the critical radius θ_c . We let $\theta_c = \pi/2$ when the threshold

radius is more than $\pi/2$. Under Assumption 3.1, we can prove that $\theta_c > 0$. Figure 1 in Kuriki and Takemura (2009) depicts an example of a tube and its critical radius.

The support cone (or tangent cone) of M at $u \in M$ is denoted by $S_u M$. (See Section 1.2 of Takemura and Kuriki (2002) for the definition.) The cone with base set M is denoted by $\text{co}(M) = \bigsqcup_{\lambda \geq 0} \lambda M$. Then, the support cone of $\text{co}(M)$ at $u \in M$ is decomposed as $S_u(\text{co}(M)) = S_u M \oplus \text{span}\{u\}$, where $\text{span}\{u\}$ is the linear space spanned by u . The normal cone of $\text{co}(M)$ at $u \in M$ is defined by the dual of the support cone: $N_u(\text{co}(M)) = S_u(\text{co}(M))^*$.

Note that the $(m-1)$ -dimensional volume of \mathbb{S}^{m-1} is $\Omega_m = 2\pi^{m/2}/\Gamma(m/2)$. For $m \times m$ matrix $A = (a_{ij})$, let $\text{tr}_0 A = 1$ and

$$\text{tr}_e A = \sum_{1 \leq k_1 < \dots < k_e \leq m} \det(a_{k_i k_j})_{1 \leq i, j \leq e}, \quad 1 \leq e \leq m$$

(Muirhead (2005), Appendix A.7). Note that $\text{tr}_1 A = \text{tr} A$, $\text{tr}_m A = \det A$. Denote the upper probability of the chi-square distribution with m degrees of freedom by $\overline{G}_m(\cdot)$. Now we can provide the upper tail probability formula for the Gaussian field $\xi^T u$, $u \in M$. The theorem below is a special case of Proposition 2.2 of Takemura and Kuriki (2002).

Proposition 3.1. *As $b \rightarrow \infty$,*

$$P\left(\max_{u \in M} \xi^T u \geq b\right) = \hat{P}(b) + O(\overline{G}_n(b^2(1 + \tan^2 \theta_c))), \quad (3.2)$$

where

$$\hat{P}(b) = \sum_{0 \leq e \leq d, e: \text{even}} w_{d+1-e} \overline{G}_{d+1-e}(b^2) + \sum_{0 \leq e \leq d-1} w'_{d-e} \overline{G}_{d-e}(b^2), \quad (3.3)$$

with

$$w_{d+1-e} = \frac{1}{\Omega_{d+1-e} \Omega_{n-d-1+e}} \int_{\text{Int} M} \left[\int_{N_u(\text{co}(M)) \cap \mathbb{S}^{n-1}} \text{tr}_e H(u, v) dv \right] du, \quad (3.4)$$

$$w'_{d-e} = \frac{1}{\Omega_{d-e} \Omega_{n-d+e}} \int_{\partial M} \left[\int_{N_u(\text{co}(M)) \cap \mathbb{S}^{n-1}} \text{tr}_e H'(u, v) dv \right] du. \quad (3.5)$$

Here, $H(u, v)$ is the second fundamental form of $\text{Int} M$ at u in the direction of v , and $H'(u, v)$ is the second fundamental form of ∂M at u in the direction of v . du is the volume element of $\text{Int} M$ or ∂M , and dv is the volume element of $N_u(\text{co}(M)) \cap \mathbb{S}^{n-1}$.

In (3.2), because of $\theta_c > 0$, the error term $O(\overline{G}_n(b^2(1 + \tan^2 \theta_c))) = O(b^{n-2} e^{-b^2(1 + \tan^2 \theta_c)/2})$ is exponentially smaller than each term $\overline{G}_j(b^2) = O(b^{j-2} e^{-b^2/2})$. Hence, (3.3) can be used as an approximation formula when b is large. The method in which $\hat{P}(b)$ is used as an approximate value is referred to as the volume-of-tube method, or simply the tube method. This name comes from the volume formula for M_θ below.

Remark 3.1. For the radius $\theta \in [0, \theta_c]$, the $(n-1)$ -dimensional spherical volume of the tube M_θ is given by

$$\text{Vol}_{n-1}(M_\theta) = \Omega_n \left\{ \sum_{0 \leq e \leq d, e: \text{even}} w_{d+1-e} \overline{B}_{\frac{1}{2}(d+1-e), \frac{1}{2}(n-d-1+e)}(\cos^2 \theta) \right. \\ \left. + \sum_{0 \leq e \leq d-1} w'_{d-e} \overline{B}_{\frac{1}{2}(d-e), \frac{1}{2}(n-d+e)}(\cos^2 \theta) \right\},$$

where w_{d+1-e} and w'_{d-e} are given in (3.4) and (3.5), $\overline{B}_{a,b}(\cdot)$ is the upper probability of the beta distribution with parameter (a, b) .

The critical radius θ_c can be evaluated using the following characterization (Theorem 4.18 of Federer (1959), Proposition 4.3 of Johansen and Johnstone (1990), Lemma 2.2 of Takemura and Kuriki (2002)). For a proof, see Theorem 2.9 of Kuriki and Takemura (2009).

Proposition 3.2. The critical radius θ_c of M is given by

$$\tan^2 \theta_c = \inf_{u \neq v \in M} \frac{(1 - u^T v)^2}{\|P_v^\perp(u - v)\|^2}, \quad (3.6)$$

where P_v^\perp is the orthogonal projection onto the normal cone $N_v(\text{co}(M))$ of $\text{co}(M)$ at v .

The local critical radius $\theta_{c, \text{loc}}$ is defined as

$$\tan^2 \theta_{c, \text{loc}} = \liminf_{u \neq v \in M, \|u-v\| \rightarrow 0} \frac{(1 - u^T v)^2}{\|P_v^\perp(u - v)\|^2}. \quad (3.7)$$

From the definition, it holds that $\theta_c \leq \theta_{c, \text{loc}}$. In general, $\theta_{c, \text{loc}}$ is easier to evaluate than θ_c .

We have summarized the volume-of-tube method to evaluate the upper tail probabilities of the maximum of random fields so far. There is another method for the same purpose referred to as the expected Euler-characteristic heuristic (Worsley (1995), Adler and Taylor (2007)). When applied to the Gaussian random field $\xi^T u$, $u \in M$, this method is stated as follows. For each b , define the excursion set by

$$A_b = \{u \in M \mid \xi^T u \geq b\}.$$

Let $\chi(\cdot)$ be the Euler-Poincaré characteristic of a set, and $\mathbf{1}(\cdot)$ be the indicator function for an event. The expected Euler-characteristic heuristic assumes that $\mathbf{1}(A_b \neq \emptyset) \approx \chi(A_b)$ for large b , and

$$P\left(\max_{u \in M} \xi^T u \geq b\right) = E[\mathbf{1}(A_b \neq \emptyset)] \approx E[\chi(A_b)].$$

Note that $\chi(A_b)$ can be evaluated by Morse's theorem, and is more tractable than $\mathbf{1}(A_b \neq \emptyset)$. Takemura and Kuriki (2002) proved the equivalence of the volume-of-tube method and the expected Euler-characteristic heuristic as follows.

Proposition 3.3 (Proposition 3.3 of Takemura and Kuriki (2002)).

$$E[\chi(A_b)] = \widehat{P}(b) \text{ for all } b \geq 0.$$

Using this, Takemura and Kuriki (2002) provided an alternative proof that the confidence band of Naiman (1986) is conservative.

4 Main results

Recall that our aim is to obtain the upper tail probability of the maximum of the Gaussian random field $Z(x, h)$ defined in (2.4), or equivalently, the chi-square random process $Y(x)$ defined in (2.5). The theorem below provides an answer. The proof is provided in the Appendix.

Theorem 4.1. *Let $\xi \sim N_n(0, I)$, $n = p(k-1)$. Let $\Gamma \subset \mathbb{S}^{p-1}$ and $M \subset \mathbb{S}^{n-1}$ be defined by (1.4) and (3.1), and let $|\Gamma|$ denote the length of Γ . Assume Assumption 3.2 on ψ . Then, as $b \rightarrow \infty$,*

$$\begin{aligned} P\left(\max_{(x,h) \in \mathcal{X} \times \mathbb{S}^{k-2}} Z(x, h) \geq b\right) &= P\left(\max_{x \in \mathcal{X}} Y(x) \geq b^2\right) \\ &= P\left(\max_{u \in M} \xi^T u \geq b\right) \\ &= \widehat{P}(b) + O(b^{n-2} e^{-(1+\tan^2 \theta_c) b^2/2}), \end{aligned}$$

where

$$\widehat{P}(b) = \frac{\Gamma(\frac{k}{2})}{\sqrt{\pi} \Gamma(\frac{k-1}{2})} |\Gamma| \{\overline{G}_k(b^2) - \overline{G}_{k-2}(b^2)\} + \chi(\Gamma) \overline{G}_{k-1}(b^2). \quad (4.1)$$

Note that if Γ (and hence M) has no boundary, then Γ is homeomorphic to \mathbb{S}^1 , and hence $\chi(\Gamma) = 0$. Otherwise, $\chi(\Gamma)$ is the number of connected components of Γ .

Theorem 4.2. *Assume Assumption 3.2. Suppose that Γ has boundaries. The approximation formula given in Theorem 4.1 is a conservative bound, that is,*

$$P\left(\max_{u \in M} \xi^T u \geq b\right) \leq \widehat{P}(b) \text{ for all } b \geq 0.$$

Proof. Arrange the $p(k-1) \times 1$ vector $\xi = (\xi_1^T, \dots, \xi_{k-1}^T)^T$, and define a $(k-1) \times p$ matrix $\Xi = (\xi_1, \dots, \xi_{k-1})^T$. Let

$$\begin{aligned} A_b &= \{u \in M \mid \xi^T u \geq b\} = \{h \otimes q \mid (q, h) \in \Gamma \times \mathbb{S}^{k-2}, h^T \Xi q \geq b\} \subset \mathbb{S}^{p(k-1)-1}, \\ \widetilde{A}_b &= \{(q, h) \in \Gamma \times \mathbb{S}^{k-2} \mid h^T \Xi q \geq b\} \subset \mathbb{S}^{p-1} \times \mathbb{S}^{k-2}, \\ B_b &= \{q \in \Gamma \mid q^T \Xi^T \Xi q \geq b^2\} \subset \mathbb{S}^{p-1}. \end{aligned}$$

Note that A_b is the excursion set of the Gaussian random field $\xi^T u$, $u \in M$, \tilde{A}_b is the excursion set of the Gaussian random field $\sum_{i=1}^{k-1} h_i(\xi_i^T q) = h^T \Xi q$, $(q, h) \in \Gamma \times \mathbb{S}^{k-2}$, and B_b is the excursion set of the chi-square random process $\sum_{i=1}^{k-1} (\xi_i^T q)^2 = q^T \Xi^T \Xi q$, $q \in \Gamma$. We will prove that for each fixed ξ , $\mathbb{1}(A_b \neq \emptyset) = \mathbb{1}(\tilde{A}_b \neq \emptyset) = \mathbb{1}(B_b \neq \emptyset)$ and $\chi(A_b) = \chi(\tilde{A}_b) = \chi(B_b)$.

First, note that because of Assumption 3.2, the map $(q, h) \mapsto h \otimes q$ is one-to-one. Hence, A_b and \tilde{A}_b are homeomorphic and therefore $\mathbb{1}(A_b \neq \emptyset) = \mathbb{1}(\tilde{A}_b \neq \emptyset)$ and $\chi(A_b) = \chi(\tilde{A}_b)$.

Moreover, noting that $\tilde{A}_b \neq \emptyset \Leftrightarrow \max_h h^T \Xi q \geq b$ for some $q \Leftrightarrow q^T \Xi \Xi^T q \geq b^2$ for some $q \Leftrightarrow B_b \neq \emptyset$, that is, $\mathbb{1}(\tilde{A}_b \neq \emptyset) = \mathbb{1}(B_b \neq \emptyset)$, we can write

$$\tilde{A}_b = \bigsqcup_{q \in B_b} \{(q, h) \mid h \in \mathbb{S}^{k-2}, h^T \Xi q \geq b\}.$$

Given $b \geq 0$, the set $\{h \in \mathbb{S}^{k-2} \mid h^T \Xi q \geq b\}$ is contractible and star-shaped about the point $h^*(q) = \Xi q / \|\Xi q\|$. That is, the map

$$\varphi : \tilde{A}_b \times [0, 1] \rightarrow \tilde{A}_b, \quad (q, h, t) \mapsto \left(q, \frac{(1-t)h + th^*(q)}{\|(1-t)h + th^*(q)\|} \right)$$

is continuous, and $\varphi(\tilde{A}_b \times \{0\}) = \tilde{A}_b$ is homotopy equivalent to the set $\varphi(\tilde{A}_b \times \{1\}) = \bigsqcup_{q \in B_b} \{(q, h^*(q))\}$. This is homotopy equivalent to $\bigsqcup_{q \in B_b} \{q\} = B_b$. Hence, $\chi(\tilde{A}_b) = \chi(B_b)$.

Recall that B_b is the excursion set of the chi-square random process on the one-dimensional index set Γ . This means that B_b is also one-dimensional, and $\chi(B_b)$ is nothing but the number of connected components of B_b . Therefore $\mathbb{1}(B_b \neq \emptyset) \leq \chi(B_b)$. By taking expectations,

$$\begin{aligned} P\left(\max_{u \in M} \xi^T u \geq b\right) &= E[\mathbb{1}(A_b \neq \emptyset)] = E[\mathbb{1}(B_b \neq \emptyset)] \\ &\leq E[\chi(B_b)] = E[\chi(A_b)] = \hat{P}(b). \end{aligned}$$

The last equality is due to Proposition 3.3. □

Remark 4.1. *Naiman (1986) proved that application of the volume-of-tube method to a Gaussian random process with a one-dimensional index set always provides a conservative band. Theorem 4.2 is a generalization of Naiman (1986)'s inequality to a chi-square random process.*

Theorem 4.3. *The interior and boundary of Γ are denoted by $\text{Int}\Gamma$ and $\partial\Gamma$, respectively. The critical radius θ_c of M is given by*

$$\tan^2 \theta_c = \min \left\{ \inf_{x \neq \tilde{x}, \psi(x) \in \text{Int}\Gamma} \frac{(1 - \alpha s)^2}{1 - s^2 - \alpha^2 t^2}, \inf_{x \neq \tilde{x}, \psi(x) \in \partial\Gamma} \frac{(1 - \alpha s)^2}{1 - s^2 - \max(0, \varepsilon(x) \alpha t)^2} \right\},$$

where the infima are taken over $x, \tilde{x} \in \mathcal{X}$, and $\alpha \in [-1, 1]$ as well as additional conditions (arguments of \inf), and

$$s = s(x, \tilde{x}) = \psi(x)^T \psi(\tilde{x}), \quad t = t(x, \tilde{x}) = \frac{\psi_x(x)^T \psi(\tilde{x})}{\|\psi_x(x)\|},$$

$$\psi_x(x) = \partial\psi(x)/\partial x,$$

$$\varepsilon(x) = \begin{cases} 1 & (\psi_x(x) \text{ is inward to } \Gamma), \\ -1 & (\psi_x(x) \text{ is outward to } \Gamma). \end{cases}$$

$\psi_x(x)$ is said to be inward or outward to Γ if the support cone of Γ at $\psi(x)$ is $S_{\psi(x)}\Gamma = \{\lambda\psi_x(s) \mid \lambda \geq 0\}$ or $\{\lambda\psi_x(s) \mid \lambda \leq 0\}$, respectively.

Theorem 4.4. Assume Assumption 3.2. Moreover, assume that $\psi : \mathcal{X} \rightarrow \mathbb{S}^{p-1}$ is of C^4 -class. Then, the local critical radius $\theta_{c,loc}$ is given by

$$\tan^2 \theta_{c,loc} = \min \left\{ \inf_{x \in \mathcal{X}: \kappa(x) \leq 2} \left(1 - \frac{\kappa(x)}{4} \right), \inf_{x \in \mathcal{X}: \kappa(x) \geq 2} \frac{1}{\kappa(x)} \right\}$$

with

$$\kappa(x) = \frac{\psi_{xx}(x)^T \psi_{xx}(x)}{(\psi_x(x)^T \psi_x(x))^2} - \frac{(\psi_{xx}(x)^T \psi_x(x))^2}{(\psi_x(x)^T \psi_x(x))^3} - 1, \quad (4.2)$$

where $\psi_x(x) = \partial\psi(x)/\partial x$ and $\psi_{xx}(x) = \partial^2\psi(x)/\partial x^2$.

The proofs of Theorems 4.3 and 4.4 are included in the Appendix.

A numerical example

At the end of this section, we provide a numerical example to discern the accuracy of the approximation formula given in Theorem 4.1, and degree of conservativeness proved by Theorem 4.2.

Suppose that $f(x) = (1, x, x^2)^T$, $\mathcal{X} = [-1, 1]$, and

$$\Sigma = \begin{pmatrix} 1 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & 1 \end{pmatrix}, \quad \Sigma^{\frac{1}{2}} = \begin{pmatrix} 1 & 0 & \frac{2}{3} \\ 0 & \sqrt{\frac{2}{3}} & 0 \\ 0 & 0 & \frac{\sqrt{5}}{3} \end{pmatrix}.$$

Then,

$$\psi(x) = \frac{1}{3(1+x^2)}(3+2x^2, \sqrt{6}x, \sqrt{5}x^2)^T, \quad |\Gamma| = \int_{\mathcal{X}} \|\dot{\psi}(x)\| dx = \int_{-1}^1 \sqrt{\frac{2}{3}} \frac{1}{1+x^2} dx = \frac{\pi}{\sqrt{6}}.$$

$\kappa(x)$ in (4.2) is always 5. Hence, the local critical radius is $\theta_{c,loc} = \tan^{-1}(1/\sqrt{5}) = 0.134\pi$. We can also confirm that the critical radius is the same as $\theta_c = \theta_{c,loc}$ using Mathematica.

Under this setting, we suppose the case of $k = 3$. The probability we need is

$$P\left(\max_{x \in [-1, 1]} Y(x) \geq b^2\right) = 1 - P(T(x, c) \leq b, \forall x \in [-1, 1], \forall c \in \mathcal{C}), \quad (4.3)$$

where

$$Y(x) = \sum_{i=1}^2 (\xi_i^T \psi(x))^2, \quad \xi_1, \xi_2 \sim N_3(0, I) \text{ i.i.d.}$$

is a chi-square random process $Y(x)$ with two degrees of freedom. The tube formula for the upper tail probability (4.3) is

$$\hat{P}(b) = \frac{\pi}{2\sqrt{6}}(\bar{G}_3(b^2) - \bar{G}_1(b^2)) + \bar{G}_2(b^2) = \left(\frac{\sqrt{\pi}}{2\sqrt{3}}b + 1\right)e^{-b^2/2}. \quad (4.4)$$

Figure 4.1 depicts the upper tail probability of the maximum (4.3) and its approximate value (4.4). We can see that the tube formula approximates the true upper tail probability sufficiently accurate in the moderate tail regions (e.g., the upper probability is less than 0.2), and that the tube formula provides a conservative bound as per Theorem 4.2.

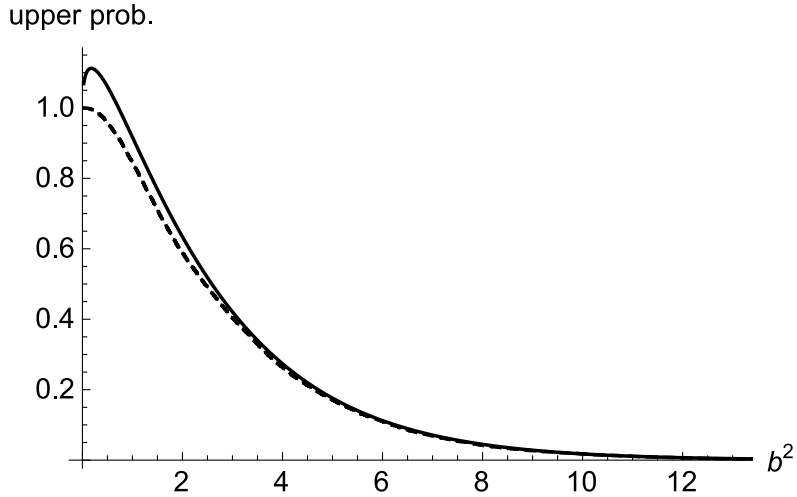


Figure 4.1: Upper tail probability of the maximum of chi-square process $Y(x)$.
(solid line: tube formula, dashed line: Monte Carlo with 10,000 replications)

We have proposed that the threshold for the confidence band should be determined as the solution $b = \hat{b}_{1-\alpha}$ for $\hat{P}(b) = \alpha$. Figure 4.2 depicts the actual confidence coefficient (coverage probability)

$$P\left(\max_{x \in [-1, 1]} Y(x) \geq \hat{b}_{1-\alpha}^2\right), \quad \alpha \in [0, 1].$$

This again shows that the confidence bands obtained by the tube method are always conservative and very accurate.

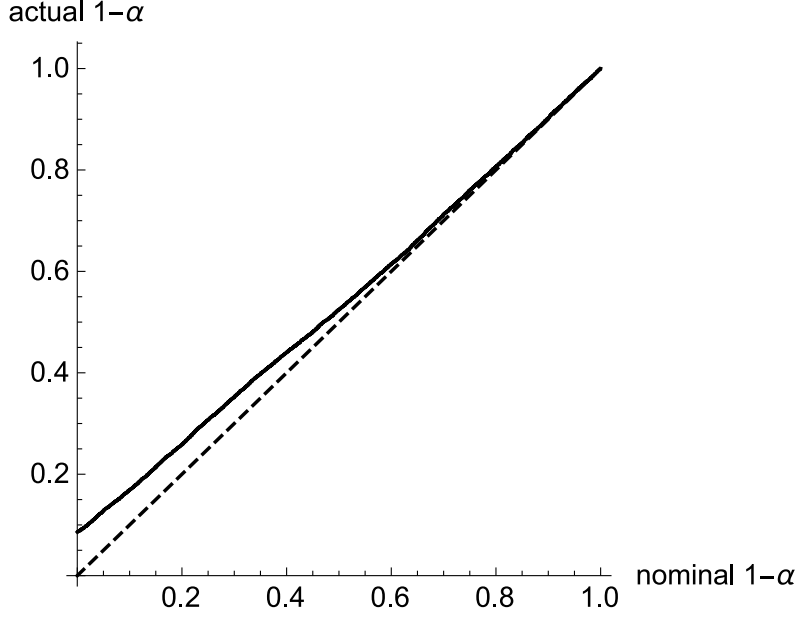


Figure 4.2: Nominal confidence coefficient vs. Actual confidence coefficient.
(solid line: actual confidence coefficient, dashed line: 45-degree line)

5 Simulation study under model misspecification

Throughout this paper, it is assumed that the nonlinear model has a finite number of basis functions $g_i(x) = \beta_i^T f(x)$ in (1.2). However, we cannot know the true model in practice. Under a slight misspecification of the model, Sun and Loader (1994) estimated the bias of the coverage probability, and proposed an adjustment to the volume-of-tube formula. Although their approach may be applied to our model, the result would be more complicated. Instead, to investigate what happens under model misspecification, we conducted a Monte Carlo simulation study in the following setting.

The domain of explanatory variable is set to be $\mathcal{X} = [0, 1]$. The data are generated from the model

$$y_{ij} = g_i(x_j) + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim N(0, 1), \quad i = 1, \dots, k, \quad j = 1, \dots, n,$$

where $k = 3$ and $n = 11$, $x_j = (j - 1)/n$, $j = 1, 2, \dots, n$. As the true regression curve is $g_i(x)$, we assume three models.

Model 1:

$$g_i(x) = \beta_i^T f_{2,5,0,1}(x), \quad \beta_1 = (0, \dots, 0)^T, \quad \beta_2 = K(0, 0, 1/2, 1, 1)^T, \quad \beta_3 = K(0, 0, 4/3, 0, 0)^T,$$

where $K = 1, 3$, or 9 ,

$$f_{d,m,a,b} = \left(B_d \left(\frac{x-a}{b-a} (m-d) - (i-d-1) \right) \right)_{i=1, \dots, m},$$

and $B_d(\cdot)$ is the B-spline function

$$B_d(x) = \sum_{r=0}^{d+1} (-1)^{d+1-r} \binom{d+1}{r} \frac{(r-x)_+^d}{d!} \quad (5.1)$$

(de Boor (1978), p. 89).

Model 2:

$$g_1(x) = 0, \quad g_2(x) = K \sin(x\pi/2), \quad g_3(x) = K \sin(x\pi), \quad K = 1, 3, 9.$$

Model 3:

$$g_1(x) = 0, \quad g_2(x) = K \frac{e^{-x/2} - e^{-x}}{e^{-1/2} - e^{-1}}, \quad g_3(x) = K \frac{\cosh(x - \frac{1}{2}) - 1}{\cosh(\frac{1}{2}) - 1}, \quad K = 1, 3, 9.$$

For all models, $g_2(x)$ is unimodal, and $g_3(x)$ is increasing. $g_2(x)$ and $g_3(x)$ are designed to have the range $[0, K]$.

To the generated data y_{ij} , we fit the curve $\beta_i^T f_{2,m,0,1}(x)$, where $m = 3, \dots, 10$. Using these models, we constructed a $1 - \alpha = 0.95$ confidence band. Coverage probabilities were estimated based on the Monte Carlo simulations with 1,000,000 replications, and are summarized in Table 5.1. In this table,

$$\delta = \max_{x \in \mathcal{X}, c \in \mathcal{C}} \left| \frac{\sum_{i=1}^k c_i ((\beta_i^*)^T f_{2,m,0,1}(x) - g_i(x))}{\sqrt{f_{2,m,0,1}(x)^T \Sigma f_{2,m,0,1}(x)}} \right|, \quad \Sigma = \left(\sum_{i=1}^n f_{2,m,0,1}(x_i)^T f_{2,m,0,1}(x_i) \right)^{-1} \quad (5.2)$$

is the bias of regression function, where β_i^* is the best parameter in the assumed model $\beta_i^T f_{2,m,0,1}(x)$.

$$\Delta = \max\{\alpha - \hat{P}(\hat{b}_{1-\alpha} + \delta), \hat{P}(\hat{b}_{1-\alpha} - \delta) - \alpha\} \quad (5.3)$$

is an approximate upper bound of the bias of coverage probability, where $\hat{b}_{1-\alpha}$ is the approximate value of $b_{1-\alpha}$ obtained by the tube method. (See A.4 for the detail.)

From this table, we first see that, for the true models ($m = 5, 8$ when model 1 is true), the coverage probabilities are more than, but approximately equal to, the nominal value 0.95, meaning that the proposed method is valid. The most remarkable point is that, throughout the study, the coverage probabilities are kept at approximately 0.95, unless the assumed model is too small, and the bias δ is large.

Table 5.2 shows the average width of the confidence band defined by

$$W = \frac{\int_{\mathcal{X}} \hat{b}_{1-\alpha} \sqrt{f(x)^T \Sigma f(x)} dx}{\int_{\mathcal{X}} dx} = \hat{b}_{0.95} \int_0^1 \sqrt{f_{2,m,0,1}(x)^T \Sigma f_{2,m,0,1}(x)} dx.$$

When the model is increasing in size, W is increasing in size. This means that a smaller model is preferable, unless it is too small to cause serious bias.

In summary, too small of a model should surely be avoided, whereas, a larger model has the disadvantage of having a wider confidence band. This trade-off is crucially important in practice, and a promising future research topic, although it is out of scope for this paper. For related topics, refer to Casella and Hwang (2012), for shrinkage confidence bands, and Leeb, et al. (2015), for confidence band post-model selection.

Table 5.1: Coverage probability under model misspecification ($1 - \alpha = 0.95$)
(prob: coverage probability, δ : bias (5.2), Δ : bound for coverage probability bias (5.3))

m	Model 1 ($K = 1$)			Model 1 ($K = 3$)			Model 1 ($K = 9$)		
	prob	δ	Δ	prob	δ	Δ	prob	δ	Δ
3	0.9365	0.4692	0.1155	0.7872	1.4076	0.8240	0.0000	4.2227	1.3542
4	0.9422	0.3996	0.0965	0.8509	1.1987	0.6909	0.0076	3.5961	1.6907
5	0.9512	0.0000	0.0000	0.9512	0.0000	0.0000	0.9512	0.0000	0.0000
6	0.9511	0.1006	0.0176	0.9477	0.3018	0.0694	0.9111	0.9053	0.4550
7	0.9514	0.0448	0.0074	0.9509	0.1343	0.0251	0.9465	0.4030	0.1100
8	0.9515	0.0000	0.0000	0.9515	0.0000	0.0000	0.9515	0.0000	0.0000
9	0.9515	0.0175	0.0029	0.9514	0.0526	0.0090	0.9504	0.1578	0.0316
10	0.9516	0.0218	0.0036	0.9515	0.0653	0.0116	0.9508	0.1959	0.0421

m	Model 2 ($K = 1$)			Model 2 ($K = 3$)			Model 2 ($K = 9$)		
	prob	δ	Δ	prob	δ	Δ	prob	δ	Δ
3	0.9509	0.06491	0.0099	0.9486	0.1947	0.0346	0.9277	0.5842	0.1640
4	0.9511	0.04999	0.0077	0.9498	0.1500	0.0264	0.9374	0.4499	0.1157
5	0.9512	0.01271	0.0019	0.9511	0.0381	0.0060	0.9504	0.1143	0.0199
6	0.9516	0.00494	0.0008	0.9516	0.0148	0.0023	0.9515	0.0445	0.0072
7	0.9514	0.00234	0.0004	0.9514	0.0070	0.0011	0.9514	0.0211	0.0034
8	0.9515	0.00137	0.0002	0.9515	0.0041	0.0007	0.9515	0.0123	0.0020
9	0.9515	0.00119	0.0002	0.9515	0.0036	0.0006	0.9515	0.0107	0.0017
10	0.9516	0.00076	0.0001	0.9516	0.0023	0.0004	0.9516	0.0068	0.0011

m	Model 3 ($K = 1$)			Model 3 ($K = 3$)			Model 3 ($K = 9$)		
	prob	δ	Δ	prob	δ	Δ	prob	δ	Δ
3	0.9512	0.02384	0.0034	0.9508	0.07152	0.0109	0.9483	0.2146	0.0390
4	0.9513	0.00766	0.0011	0.9512	0.02299	0.0034	0.9511	0.0690	0.0109
5	0.9512	0.00218	0.0003	0.9512	0.00653	0.0010	0.9512	0.0196	0.0030
6	0.9516	0.00095	0.0002	0.9516	0.00285	0.0004	0.9516	0.0086	0.0013
7	0.9514	0.00046	0.0001	0.9514	0.00138	0.0002	0.9514	0.0042	0.0006
8	0.9515	0.00028	0.0000	0.9514	0.00083	0.0001	0.9515	0.0025	0.0004
9	0.9515	0.00024	0.0000	0.9515	0.00071	0.0001	0.9515	0.0021	0.0003
10	0.9516	0.00014	0.0000	0.9516	0.00042	0.0001	0.9516	0.0013	0.0002

Table 5.2: Average band-width W ($1 - \alpha = 0.95$)

m	3	4	5	6	7	8	9	10
W	1.463	1.752	2.017	2.275	2.546	2.764	2.990	3.211

6 Growth curve analysis

As mentioned in Section 1, the analysis of growth curves is one of our research objectives to which we apply our method. In this section, we demonstrate the analysis of mouse growth as an illustration. Sun, Raz, and Faraway (1999) proposed simultaneous confidence bands for a growth curve by virtue of the volume-of-tube method. Differently from their analysis, we focus on the contrast of several growth curves.

Mice are one of the most popular model organisms, and are often used in genomic research. Figure 6.1 depicts the average body weights of male mice from four different strains measured from 2 to 20 weeks after birth. The four strains are C57BL/6 (referred to as B6), MSM/Ms (MSM), B6-Chr17^{MSM}(B6-17), and B6-ChrXT^{MSM}(B6-XT). Among these, B6 is the most common laboratory strain and serves as the standard. MSM is a wild-derived strain that has contrasting properties to B6 such as non-black color, small size, and aggressive behavior. B6-17 and B6-XT are artificial strains known as consomic mice made from B6 and MSM. B6-17 has all the chromosomes from B6, but only chromosome 17 from MSM; B6-XT has all the chromosomes from B6, but only half of the X chromosome from MSM. By comparing the consomic strains with B6, we expect to reveal the role of each chromosome.

The dataset we used is publicly available as Supplemental Table S1 of Takada, et al. (2008). In their experiments, the weight (unit: gram) y_{ijh} of the h th individual from strain i was measured at time point x_j . The measurement time points were $\{x_1, \dots, x_{10}\} = \{2, 4, \dots, 20\}$ ($n = 10$). This dataset includes the average body weight y_{ij} of strain i at time x_j , and its standard error

$$y_{ij} = \frac{1}{r_i} \sum_{h=1}^{r_i} y_{ijh}, \quad \widehat{\text{s.e.}}(y_{ij}) = \sqrt{\frac{1}{r_i^2} \sum_{h=1}^{r_i} (y_{ijh} - y_{ij})^2},$$

as well as the number r_i of individuals of strain i .

In the following analysis, we use $k = 3$ groups (strains) B6 ($i = 1$), B6-17 ($i = 2$), and B6-XT ($i = 3$). The number of individuals are $r_1 = 12$, $r_2 = 24$, and $r_3 = 12$.

We fit the model (1.1) to these data. We estimate the variance as

$$\widehat{\sigma}(x_j)^2 = \frac{1}{\sum_{i=1}^k (r_i - 1)} \sum_{i=1}^k \sum_{h=1}^{r_i} (y_{ijh} - y_{ij})^2 = \frac{1}{\sum_{i=1}^k (r_i - 1)} \sum_{i=1}^k r_i^2 \widehat{\text{s.e.}}(y_{ij})^2,$$

which is used as the true value $\sigma(x_j)^2$ hereafter. Figure 6.2 plots the estimated standard

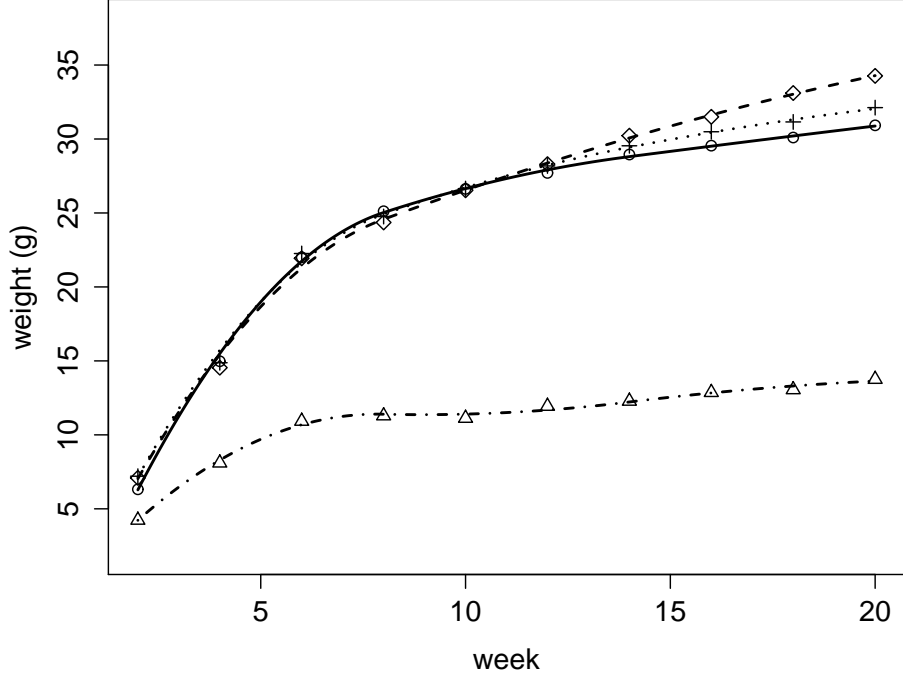


Figure 6.1: Average body weights of mice from four strains.

(sample mean: \circ (B6), $+$ (B6-17), \diamond (B6-XT), \triangle (MSM);
fitted curve: — (B6), \cdots (B6-17), $--$ (B6-XT), $-\cdot-$ (MSM))

error $\hat{\sigma}(x_j)$. One particular feature of this dataset is that the experiment is well controlled and the measurement errors are quite small.

As the basis function $f(x)$, we consider a family of basis functions

$$f(x) = f_{d,m,2,20}(x) = \left(B_d \left(\frac{x-2}{20-2} (m-d) - (i-d-1) \right) \right)_{1 \leq i \leq m},$$

with $B_d(x)$ given in (5.1). $f_{d,m,2,20}(x)$ consists of m B-spline bases with equally-spaced knots at intervals of $(20-2)/(m-d)$. Note that $f_{d,m,2,20}(x)$ is piecewise of class C^d .

In the range $d = 2, 3, 4$ and $m = d+1, d+2, \dots, n (= 10)$, we searched for the best model that minimizes AIC and BIC defined below:

$$\text{AIC}_{d,m} = L_{d,m} + 2km, \quad \text{BIC}_{d,m} = L_{d,m} + \sum_{i=1}^k \log(r_i n) m, \quad L_{d,m} = \sum_{i=1}^k r_i \sum_{j=1}^n \frac{(y_{ij} - \hat{y}_{ij})^2}{\sigma(x_j)^2}$$

with $k = 3$, $n = 10$, where $\hat{y}_{ij} = \hat{\beta}_i^T f_{d,m,2,20}(x_j)$. In both criteria, the minimizer was $(d, m) = (2, 5)$, which we use as the true value hereafter.

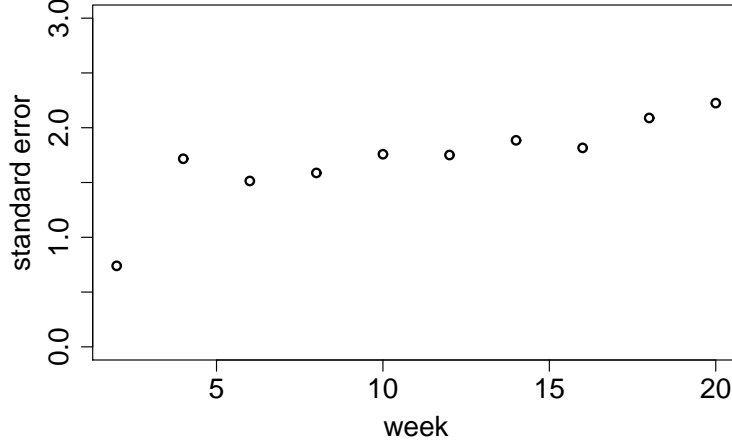


Figure 6.2: Estimated standard error $\hat{\sigma}(x_j)$.

Suppose that we are interested in the period $\mathcal{X} = [a, b] = [2, 20]$. An approximate value of the length of Γ in (1.4) is given by

$$|\Gamma| \approx \sum_{t=1}^N \|\psi(x_t) - \psi(x_{t-1})\|,$$

where $x_t = a + t(b - a)/N$, $t = 0, 1, \dots, N$. When $N = 10,000$, the approximate value of $|\Gamma|$ is $6.989 = 2.225\pi$. Using this value, the critical value is $b_{1-\alpha} = 3.258$ ($\alpha = 0.05$).

To compare k groups, various types of contrasts are used. For a pairwise comparison between group i and group j , we choose $c = (\dots, 0, \frac{1}{i\text{th}}, 0, \dots, 0, -\frac{1}{j\text{th}}, 0, \dots)$. For the comparison of groups $\{i, j\}$ and group k , we use

$$c = \left(\dots, 0, \frac{r_i}{r_i + r_j}, 0, \dots, 0, \frac{r_j}{r_i + r_j}, 0, \dots, 0, -1, 0, \dots \right).$$

Figure 6.3 depicts the difference curves of strains B6-17 vs. B6 (left) and B6-XT vs. B6 (right), and their 95% simultaneous confidence bands. In the left panel, the horizontal line representing zero difference is almost between the confidence bands. This means that there is no significant difference between B6-17 and B6. In contrast, in the right panel, after about week 14, the horizontal line is outside the confidence bands, thereby indicating that B6-XT and B6 are different during this period.

For a fixed x , the test statistic for the null hypothesis $H_{0,x} : \beta_1^T f(x) = \dots = \beta_k^T f(x)$ is

$$\chi^2(x) = \frac{1}{f(x)^T \Sigma f(x)} \sum_{i=1}^k r_i \left(\hat{\beta}_i^T f(x) - \frac{\sum_{i=1}^k r_i \hat{\beta}_i^T f(x)}{\sum_{i=1}^k r_i} \right)^2.$$

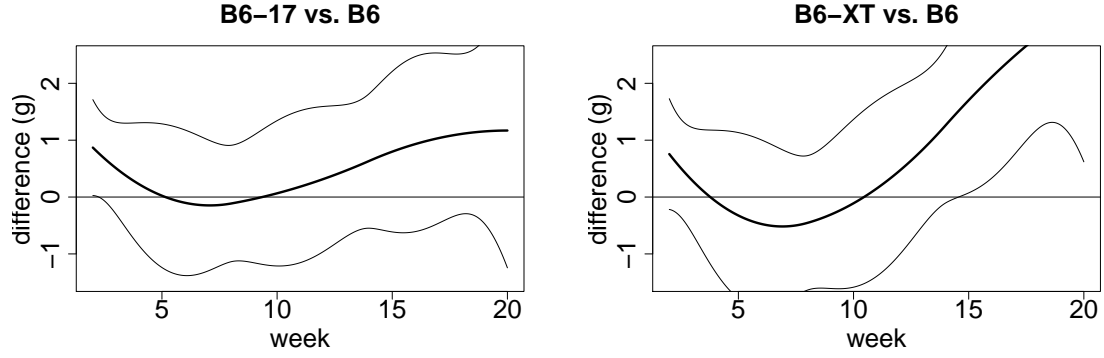


Figure 6.3: Differences of body weights and 95% confidence bands.

For a fixed x , the null distribution is the chi-square distribution with $k - 1$ degrees of freedom. However, for the overall null hypothesis $H_0 : \beta_1^T f(x) = \dots = \beta_k^T f(x)$ for all $x \in \mathcal{X}$, the distribution of the maximum of the chi-square random process should be used. Figure 6.4 shows $\chi^2(x)$ and its upper 5% critical value $b_{0.95}^2$. As already shown in Figure 6.3, after about week 14, the hypothesis of equality is rejected.

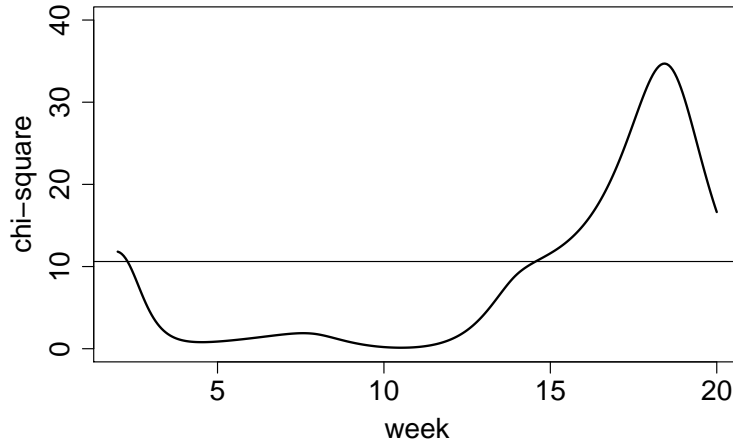


Figure 6.4: Chi-square process $\chi^2(x)$ and its upper 5% critical value.

A Appendix: Proofs

A.1 Proof of Theorem 4.1

Contribution of the inner points $\text{Int } M$

We obtain here the coefficients w_{d+1-e} in (3.4) when M is given in (3.1).

Let $h = h(\theta)$, $\theta = (\theta_i)_{1 \leq k-2}$, be a local coordinate system of \mathbb{S}^{k-2} . For example,

$$h = h(\theta) = \begin{pmatrix} \cos \theta_1 \\ \sin \theta_1 \cos \theta_2 \\ \sin \theta_1 \sin \theta_2 \cos \theta_3 \\ \vdots \\ \sin \theta_1 \cdots \sin \theta_{k-3} \cos \theta_{k-2} \\ \sin \theta_1 \cdots \sin \theta_{k-3} \sin \theta_{k-2} \end{pmatrix}_{(k-1) \times 1},$$

where

$$\theta \in \Theta = \{(\theta_1, \dots, \theta_{k-2}) \mid 0 \leq \theta_i \leq \pi \ (i = 1, \dots, k-3), \ 0 \leq \theta_{k-2} < 2\pi\}.$$

Let $(x, \theta) \in \mathcal{X} \times \Theta$ be fixed, and let $\phi(x, \theta) = h(\theta) \otimes \psi(x) \in M$. We write $\psi = \psi(x)$, $h = h(\theta)$ and $\phi = \phi(x, \theta)$ for simplicity. We first assume that $x \in \text{Int } \mathcal{X}$, and hence $\phi(x, \theta) \in \text{Int } M$.

By applying the Gram-Schmidt orthonormalization to the sequence $\psi, \partial\psi/\partial x, \partial^2\psi/\partial x^2, \dots$, we construct the ONB (orthonormal basis) $\psi_{(i)}$, $i = 0, \dots, p-1$, of \mathbb{R}^p . The first three bases are

$$\psi_{(0)} = \psi, \quad \psi_{(1)} = \frac{1}{\sqrt{g}} \frac{\partial\psi}{\partial x}, \quad \psi_{(2)} = \frac{1}{\sqrt{\eta - \frac{\gamma^2}{g} - g^2}} \left(\frac{\partial^2\psi}{\partial x^2} + g\psi - \frac{\gamma}{g} \frac{\partial\psi}{\partial x} \right),$$

where

$$g = g(x) = \left(\frac{\partial\psi}{\partial x} \right)^T \left(\frac{\partial\psi}{\partial x} \right), \quad \gamma = \gamma(x) = \left(\frac{\partial^2\psi}{\partial x^2} \right)^T \left(\frac{\partial\psi}{\partial x} \right), \quad \eta = \eta(x) = \left(\frac{\partial^2\psi}{\partial x^2} \right)^T \left(\frac{\partial^2\psi}{\partial x^2} \right).$$

Similarly, from the sequence $h, \partial h/\partial \theta_i$, $i = 1, \dots, k-2$, we obtain the ONB $h_{(i)}$, $i = 0, \dots, k-2$, of \mathbb{R}^{k-1} . We prepare a $(k-2) \times (k-2)$ upper triangle matrix D such that

$$\left(h, \frac{\partial h}{\partial \theta_1}, \dots, \frac{\partial h}{\partial \theta_{k-2}} \right) = (h_{(0)}, h_{(1)}, \dots, h_{(k-2)}) \begin{pmatrix} 1 & 0 \\ 0 & D \end{pmatrix}, \quad \text{or} \quad D = \left(h_{(i)}^T \frac{\partial h}{\partial \theta_j} \right)_{1 \leq i, j \leq k-2}.$$

Now we have the ONB $h_{(i)} \otimes \psi_{(j)}$, $i = 0, \dots, k-2$, $j = 0, \dots, p-1$, of the ambient space \mathbb{R}^n with $n = p(k-1)$. Note that $\phi = h_{(0)} \otimes \psi_{(0)}$.

The tangent space $T_\phi M$ is spanned by

$$\frac{\partial\phi}{\partial x} = h \otimes \frac{\partial\psi}{\partial x}, \quad \frac{\partial\phi}{\partial \theta_i} = \frac{\partial h}{\partial \theta_i} \otimes \psi, \quad i = 1, \dots, k-2.$$

The metric matrix of $T_\phi M$ with respect to the parameter $x, \theta_1, \dots, \theta_{k-2}$ is

$$\begin{pmatrix} g & 0 \\ 0 & G \end{pmatrix}, \quad \text{where } G = \left(\left(\frac{\partial h}{\partial \theta_i} \right)^T \left(\frac{\partial h}{\partial \theta_j} \right) \right)_{1 \leq i, j \leq k-2} = D^T D. \quad (\text{A.1})$$

$T_\phi M$ has the ONB $h_{(0)} \otimes \psi_{(1)}, h_{(i)} \otimes \psi_{(0)}, i = 1, \dots, k-2$.

The normal space perpendicular to $T_\phi(\text{co}(M)) = T_\phi M \oplus \text{span}\{\phi\}$ is

$$N_\phi(\text{co}(M)) = \text{span}\{h_{(0)} \otimes \psi_{(j)}, j = 2, \dots, p-1; \\ h_{(i)} \otimes \psi_{(j)}, i = 1, \dots, k-2, j = 1, \dots, p-1\}. \quad (\text{A.2})$$

The second order derivatives of $\phi = \phi(x, \theta)$ are

$$\frac{\partial^2 \phi}{\partial x^2} = h \otimes \frac{\partial^2 \psi}{\partial x^2}, \quad \frac{\partial^2 \phi}{\partial x \partial \theta_i} = \frac{\partial h}{\partial \theta_i} \otimes \frac{\partial \psi}{\partial x}, \quad \frac{\partial^2 \phi}{\partial \theta_i \partial \theta_j} = \frac{\partial^2 h}{\partial \theta_i \partial \theta_j} \otimes \psi.$$

Taking the inner product of the second derivatives and the ONB of $N_\phi(\text{co}(M))$ listed in (A.2), we see that the nonzero elements of the second fundamental form are

$$-\left(h \otimes \frac{\partial^2 \psi}{\partial x^2}\right)^T (h_{(0)} \otimes \psi_{(2)}) = -\left(\frac{\partial^2 \psi}{\partial x^2}\right)^T \psi_{(2)} = -\zeta,$$

where

$$\zeta = \zeta(x) = \sqrt{\eta - \frac{\gamma^2}{g} - g^2}$$

and

$$-\left(\frac{\partial h}{\partial \theta_i} \otimes \frac{\partial \psi}{\partial x}\right)^T (h_{(j)} \otimes \psi_{(1)}) = -\left(\frac{\partial h}{\partial \theta_i}\right)^T h_{(j)} \sqrt{g} = -D_{ji} \sqrt{g}.$$

We renumber the ONB of $N_\phi(\text{co}(M))$ as

$$N_1 = h_{(0)} \otimes \psi_{(2)}, \quad N_i = h_{(i-1)} \otimes \psi_{(1)}, i = 2, \dots, k-1,$$

and $N_k, N_{k+1}, \dots, N_{pk-p-k}$ are the other vectors. Write $N(t) = \sum_{i=1}^{pk-p-k} N_i t_i$, where $t = (t_1, \dots, t_{pk-p-k})$. Then,

$$\begin{aligned} -\left(\frac{\partial^2 \phi}{\partial x^2}\right)^T N(t) &= -\zeta t_1, \\ -\left(\frac{\partial^2 \phi}{\partial x \partial \theta_i}\right)^T N(t) &= -\sum_{j=1}^{k-2} D_{ji} t_{j+1} \sqrt{g}, \\ -\left(\frac{\partial^2 \phi}{\partial \theta_i \partial \theta_j}\right)^T N(t) &= 0. \end{aligned}$$

Therefore, the second fundamental form (unnormalized version) in the direction $N(t)$ is

$$\begin{pmatrix} -\zeta t_1 & -(t_2, \dots, t_{k-1})D\sqrt{g} \\ -D^T \begin{pmatrix} t_2 \\ \vdots \\ t_{k-1} \end{pmatrix} \sqrt{g} & 0 \end{pmatrix}. \quad (\text{A.3})$$

Multiplication of the inverse of the metric (A.1) enables us to obtain the normalized version of the second fundamental form. Noting that

$$\begin{pmatrix} g & 0 \\ 0 & G \end{pmatrix}^{-1} = \begin{pmatrix} g & 0 \\ 0 & D^T D \end{pmatrix}^{-1} = \begin{pmatrix} 1/\sqrt{g} & 0 \\ 0 & D^{-1} \end{pmatrix} \begin{pmatrix} 1/\sqrt{g} & 0 \\ 0 & (D^T)^{-1} \end{pmatrix},$$

we multiply

$$\begin{pmatrix} 1/\sqrt{g} & 0 \\ 0 & (D^T)^{-1} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1/\sqrt{g} & 0 \\ 0 & D^{-1} \end{pmatrix}$$

from the left and the right to (A.3), respectively, to obtain

$$\begin{aligned} & \begin{pmatrix} 1/\sqrt{g} & 0 \\ 0 & (D^T)^{-1} \end{pmatrix} \begin{pmatrix} -\zeta t_1 & -(t_2, \dots, t_{k-1})D\sqrt{g} \\ -D^T \begin{pmatrix} t_2 \\ \vdots \\ t_{k-1} \end{pmatrix} \sqrt{g} & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{g} & 0 \\ 0 & D^{-1} \end{pmatrix} \\ &= \begin{pmatrix} -(\zeta/g)t_1 & -(t_2, \dots, t_{k-1}) \\ -\begin{pmatrix} t_2 \\ \vdots \\ t_{k-1} \end{pmatrix} & 0 \end{pmatrix} = H(x, \theta; N(t)). \end{aligned}$$

This is the second fundamental form with respect to the orthonormal coordinates. Now we have

$$\text{tr}_e H(x, \theta; N(t)) = \begin{cases} 1 & (e = 0), \\ -(\zeta/g)t_1 & (e = 1), \\ -\sum_{j=2}^{k-1} t_j^2 & (e = 2), \\ 0 & (\text{otherwise}). \end{cases} \quad (\text{A.4})$$

Next, we evaluate the integral

$$\int_{v \in N_\phi(\text{co}(M)) \cap \mathbb{S}^{n-1}} \text{tr}_e H(x, \theta; v) dv, \quad (\text{A.5})$$

where $n = p(k-1)$, dv is the volume element of $N_\phi(\text{co}(M)) \cap \mathbb{S}^{n-1}$, by following Kuriki and Takemura (2001, Section 4.2.2). Recall that $d = \dim M = k-1$.

Because $N_\phi(\text{co}(M))$ is a linear space of dimension $n - d - 1 = p(k - 1) - (k - 1) - 1 = pk - p - k$, $N_\phi(\text{co}(M)) \cap \mathbb{S}^{n-1}$ is nothing but a $(pk - p - k - 1)$ -dimensional unit sphere. Hence,

$$\int_{v \in N_\phi(\text{co}(M)) \cap \mathbb{S}^{n-1}} dv = \text{Vol}(\mathbb{S}^{pk-p-k-1}) = \Omega_{pk-p-k}.$$

Therefore, if $V \sim \text{Unif}(N_\phi M \cap \mathbb{S}^{n-1})$, then (A.5) = $\Omega_{pk-p-k} \times E[\text{tr}_e H(x, \theta; V)]$.

Suppose that $T = (T_1, \dots, T_{pk-p-k}) \sim N_{pk-p-k}(0, I)$, and let $N(T) = \sum_{i=1}^{pk-p-k} N_i T_i$. Then,

$$\|N(T)\|^2 = \sum_{i=1}^{pk-p-k} T_i^2 \sim \chi_{pk-p-k}^2 \quad \text{and} \quad V = \frac{N(T)}{\|N(T)\|} \sim \text{Unif}(N_\phi M \cap \mathbb{S}^{n-1})$$

are independently distributed. Hence,

$$E[\text{tr}_e H(x, \theta; N(T))] = E[\|N(T)\|^e \text{tr}_e H(x, \theta, V)] = E[\|N(T)\|^e] E[\text{tr}_e H(x, \theta, V)],$$

and

$$E[\text{tr}_e H(x, \theta, V)] = \frac{E[\text{tr}_e H(x, \theta; N(T))]}{E[(\chi_{pk-p-k}^2)^{e/2}]}.$$

From (A.4),

$$E[\text{tr}_e H(x, \theta; N(T))] = \begin{cases} 1 & (e = 0), \\ 0 & (e = 1), \\ -\sum_{i=2}^{k-1} E[T_i^2] = -(k-2) & (e = 2), \\ 0 & (\text{otherwise}), \end{cases}$$

and hence,

$$E[\text{tr}_e H(x, \theta; V)] = \begin{cases} 1 & (e = 0), \\ -\frac{k-2}{pk-p-k} & (e = 2), \\ 0 & (\text{otherwise}). \end{cases}$$

Therefore,

$$\int_{v \in N_\phi(\text{co}(M)) \cap \mathbb{S}^{n-1}} \text{tr}_e H(x, \theta, v) dv = \begin{cases} \Omega_{pk-p-k} & (e = 0), \\ -\frac{k-2}{pk-p-k} \Omega_{pk-p-k} & (e = 2), \\ 0 & (\text{otherwise}). \end{cases}$$

Note that the results are independent of x and θ . This implies that the integral with respect to the volume element du in (3.4) of $\text{Int}M$ at $\phi(x, \theta)$, that is, the integral with respect to

$$du = \sqrt{g} dx \times |G(\theta)|^{\frac{1}{2}} d\theta_1 \cdots d\theta_{k-2},$$

is simply multiplying the constant

$$\text{Vol}(M) = |\Gamma| \times \text{Vol}(\mathbb{S}^{k-2}) = |\Gamma| \Omega_{k-1}.$$

Finally, from (3.4) of Proposition 3.1,

$$\begin{aligned} w_{d+1} = w_k &= \frac{1}{\Omega_k \Omega_{pk-p-k}} \times |\Gamma| \Omega_{k-1} \Omega_{pk-p-k}, \\ w_{d-1} = w_{k-2} &= \frac{1}{\Omega_{k-2} \Omega_{pk-p-k+2}} \times |\Gamma| \Omega_{k-1} \times \left(-\frac{k-2}{pk-p-k} \right) \Omega_{pk-p-k}, \end{aligned}$$

and the other w 's are zero. Simple calculations give

$$w_k = -w_{k-2} = \frac{\Gamma(\frac{k}{2})}{\sqrt{\pi} \Gamma(\frac{k-1}{2})} |\Gamma|. \quad (\text{A.6})$$

Contribution of the boundary ∂M

We obtain here the coefficients w'_{d-e} in (3.5) when M is given in (3.1).

Suppose that $\mathcal{X} = [a, b]$. Then,

$$\partial M = \{\phi(a, \theta) \mid \theta \in \Theta\} \sqcup \{\phi(b, \theta) \mid \theta \in \Theta\}.$$

Let $x = a$ and $\theta \in \Theta$ be fixed. Then, $\phi(a, \theta) \in \partial M$. The metric of the boundary ∂M at $\phi(x, \theta)$ is

$$\left(\frac{\partial \phi}{\partial \theta_i} \right)^T \left(\frac{\partial \phi}{\partial \theta_j} \right) \Big|_{(a, \theta)} = (G(\theta))_{ij}.$$

Note that $\partial(\text{co}(M)) = \text{co}(\partial M)$. The support cone of $\text{co}(M)$ at $\phi(a, \theta) \in \partial(\text{co}(M))$ is

$$S_{\phi(a, \theta)}(\text{co}(M)) = L \oplus K_1,$$

where

$$L = \text{span} \left\{ h \otimes \psi, \frac{\partial h}{\partial \theta_i} \otimes \psi, i = 1, \dots, k-2 \right\}, \quad K_1 = \left\{ \lambda \left(h \otimes \frac{\partial \psi}{\partial x} \right) \mid \lambda \geq 0 \right\}.$$

This is a direct sum (the Minkowski sum) of a linear subspace and a cone. To obtain its dual cone, the following lemma is useful.

Lemma. *Let K_1 be a cone, and let L be a linear subspace. Let $K = K_1 \oplus L$. Then, $K^* = K_1^* \cap L^\perp$.*

Because

$$K_1^* = \{ \lambda (h_{(0)} \otimes \psi_{(1)}) \mid \lambda \leq 0 \} \oplus \text{span} \{ h_{(0)} \otimes \psi_{(1)} \}^\perp$$

and

$$L^\perp = \text{span} \{ h_{(i)} \otimes \psi_{(j)}, i = 0, \dots, k-2, j = 1, \dots, p-1 \},$$

we have

$$N_{\phi(a,\theta)}(\text{co}(M)) = K^* = K_1^* \cap L^\perp = \{ \lambda(h_{(0)} \otimes \psi_{(1)}) \mid \lambda \leq 0 \} \oplus \text{span} \{ \\ h_{(0)} \otimes \psi_{(j)}, j = 2, \dots, p-1; \\ h_{(i)} \otimes \psi_{(j)}, i = 1, \dots, k-2, j = 1, \dots, p-1 \},$$

with $\dim N_{\phi(a,\theta)}(\text{co}(M)) = pk - p - k + 1$.

The second fundamental form of $\text{co}(\partial M)$ at $\phi(a, \theta)$ is

$$-\left(\frac{\partial^2 \phi(a, \theta)}{\partial \theta_i \partial \theta_j}\right)^T v = -\left(\frac{\partial^2 h}{\partial \theta_i \partial \theta_j} \otimes \psi\right)^T v, \quad v \in N_{\phi(a,\theta)}(\text{co}(M)). \quad (\text{A.7})$$

We can easily see that the second fundamental form (A.7) is always zero. Therefore, the contribution of the boundary to w'_{d-e} in (3.5) is only to case $e = 0$. That is, all w'_i but w'_{k-1} are zero. The contribution of the boundary $\{\phi(a, \theta) \mid \theta \in \Theta\}$ to w'_{k-1} is

$$\frac{1}{\Omega_{k-1-0}\Omega_{p(k-1)-(k-1)+0}} \int_{\partial M} |G|^{\frac{1}{2}} d\theta \int_{N_{\phi(a,\theta)}(\text{co}(M)) \cap \mathbb{S}^{n-1}} dv = \frac{\text{Vol}(\mathbb{S}^{k-2}) \text{Vol}(\text{half of } \mathbb{S}^{pk-p-k})}{\Omega_{k-1}\Omega_{pk-p-k+1}} \\ = \frac{1}{2}.$$

The contribution of the other boundary $\{\phi(b, \theta) \mid \theta \in \Theta\}$ to w_{k-1} has the same value of $1/2$. Moreover, if the number of connected components of Γ exceeds one, we need to select all boundaries. Since the number of boundaries is $2\chi(\Gamma)$, we have

$$w'_{k-1} = \frac{1}{2} \times 2\chi(\Gamma) = \chi(\Gamma). \quad (\text{A.8})$$

Substituting (A.6) and (A.8) into (3.3) yields (4.1).

A.2 Proof of Theorem 4.3

We apply the formula (3.6) for θ_c to the case where M is given in (3.1).

Let

$$u = \phi(\tilde{x}, \tilde{\theta}) = h(\tilde{\theta}) \otimes \psi(\tilde{x}), \quad v = \phi(x, \theta) = h(\theta) \otimes \psi(x),$$

and write $h = h(\theta)$, $\tilde{h} = h(\tilde{\theta})$, $\psi = \psi(x)$, $\tilde{\psi} = \psi(\tilde{x})$. We discuss the two cases (i) $\psi \in \text{Int}\Gamma$ and (ii) $\psi \in \partial\Gamma$ separately.

Case (i). Suppose that $\psi(x) \in \text{Int}\Gamma$. Write $\phi_{\theta_i} = (\partial h(\theta)/\partial \theta_i) \otimes \psi(x)$, $\phi_x = h(\theta) \otimes \psi_x(x)$, $\psi_x = \partial \psi(x)/\partial x$.

The orthogonal projection matrix onto the space $T_\phi(\text{co}(M)) = \text{span}\{\phi, \phi_{\theta_i}, \phi_x\}$ is

$$\begin{aligned}
P_v &= (\phi \quad \phi_{\theta_i} \quad \phi_x)_{p(k-1) \times k} \left(\begin{pmatrix} \phi^T \\ \phi_{\theta_i}^T \\ \phi_x^T \end{pmatrix} (\phi \quad \phi_{\theta_i} \quad \phi_x) \right)_{k \times k}^{-1} \begin{pmatrix} \phi^T \\ \phi_{\theta_i}^T \\ \phi_x^T \end{pmatrix}_{k \times p(k-1)} \\
&= (\phi \quad \phi_{\theta_i} \quad \phi_x) \begin{pmatrix} 1 & 0 & 0 \\ 0 & G(\theta) & 0 \\ 0 & 0 & g(x) \end{pmatrix}^{-1} \begin{pmatrix} \phi^T \\ \phi_{\theta_i}^T \\ \phi_x^T \end{pmatrix} \\
&= \phi\phi^T + (\phi_{\theta_i})G(\theta)^{-1}(\phi_{\theta_i})^T + \frac{1}{g(x)}\phi_x\phi_x^T \\
&= hh^T \otimes \psi\psi^T + (I_{k-1} - hh^T) \otimes \psi\psi^T + \frac{1}{g}hh^T \otimes \psi_x\psi_x^T \\
&= I_{k-1} \otimes \psi\psi^T + \frac{1}{g}hh^T \otimes \psi_x\psi_x^T.
\end{aligned}$$

As $P_v^\perp(w) = I(w) - P_v(w)$,

$$\frac{(1 - u^T v)^2}{\|(I_n - P_v)u\|^2} = \frac{(1 - (\tilde{h} \otimes \tilde{\psi})^T(h \otimes \psi))^2}{\|(I_{k-1} \otimes I_p - I_{k-1} \otimes \psi\psi^T - \frac{1}{g}hh^T \otimes \psi_x\psi_x^T)(\tilde{h} \otimes \tilde{\psi})\|^2}. \quad (\text{A.9})$$

Let $s = \psi^T(x)\psi(\tilde{x}) = \psi^T\tilde{\psi}$, $r = \psi_x^T(x)\psi(\tilde{x}) = \psi_x^T\tilde{\psi}$, $\alpha = h(\theta)^T h(\tilde{\theta}) = h^T\tilde{h}$. The numerator in (A.9) is $(1 - \alpha s)^2$. The denominator in (A.9) is

$$\begin{aligned}
\left\| \tilde{h} \otimes \tilde{\psi} - s\tilde{h} \otimes \psi - \frac{\alpha r}{g}h \otimes \psi_x \right\|^2 &= 1 + s^2 + \frac{\alpha^2 r^2}{g} - 2s^2 - 2\frac{\alpha r}{g}\alpha r \\
&= 1 - s^2 - \frac{\alpha^2 r^2}{g} = 1 - s^2 - \alpha^2 t^2,
\end{aligned}$$

where $t = r/\sqrt{g} = \psi_x(x)^T \psi(\tilde{x})/\|\psi_x(x)\|$. Hence,

$$(\text{A.9}) = \frac{(1 - \alpha s)^2}{1 - s^2 - \alpha^2 t^2}.$$

The infimum is taken over “ $x \neq \tilde{x}$ or $\theta \neq \tilde{\theta}$ ”, or equivalently, “ $x \neq \tilde{x}$ or $\alpha \neq 1$ ”. However, when $x = \tilde{x}$ and $\alpha \neq 1$, the argument of the infimum is $(1 - \alpha)^2/0 = \infty$. Therefore, we can exclude this case $x = \tilde{x}$ from the infimum argument.

Case (ii). Suppose that $\psi(x) \in \partial\Gamma$. Fix a point on the boundary

$$v = \phi(x, \theta) = h(\theta) \otimes \psi(x) \in \partial M.$$

The support cone of $\text{co}(M)$ at v is

$$S_v(\text{co}(M)) = \text{span}\{\phi, \phi_{\theta_i}\} \oplus \{\lambda \varepsilon \phi_x \mid \lambda \geq 0\}$$

where $\varepsilon = \varepsilon(x) = 1$ if ψ_x is inward to Γ , $= -1$ if ψ_x is outward to Γ .

The orthogonal projection operator onto the cone $S_v(\text{co}(M))$ is $w \mapsto P_v(w)$, where

$$\begin{aligned} P_v(w) &= \phi\phi^T w + \phi_\theta G^{-1}(\theta)\phi_\theta^{-1} w + \frac{\phi_x}{\|\phi_x\|^2} \max\{0, \varepsilon\phi_x^T w\} \\ &= (I_{k-1} \otimes \psi\psi^T)w + \frac{\phi_x}{\|\phi_x\|^2} \max\{0, \varepsilon\phi_x^T w\}. \end{aligned}$$

Hence,

$$P_v^\perp(w) = w - P_v(w) = w - (I_k \otimes \psi\psi^T)w - \frac{h \otimes \psi_x}{g} \max\{0, \varepsilon(h \otimes \psi_x)^T w\}.$$

Substituting $u = \phi(\tilde{x}, \tilde{\theta}) = \tilde{h} \otimes \tilde{\psi}$,

$$P_v^\perp(u - v) = \tilde{h} \otimes \tilde{\psi} - s\tilde{h} \otimes \psi - \frac{\max\{0, \varepsilon\alpha r\}}{g} h \otimes \psi_x,$$

$$\begin{aligned} \|P_v^\perp(u - v)\|^2 &= 1 + s^2 + \frac{\max\{0, \varepsilon\alpha r\}^2}{g} - 2s^2 - 2\alpha r \frac{\max\{0, \varepsilon\alpha r\}}{g} \\ &= 1 - s^2 - \frac{\max\{0, \varepsilon\alpha r\}^2}{g} = 1 - s^2 - \max\{0, \varepsilon\alpha t\}^2, \end{aligned}$$

and we have

$$\frac{(1 - u^T v)^2}{\|P_v^\perp(u - v)\|^2} = \frac{(1 - \alpha s)^2}{1 - s^2 - \max\{0, \varepsilon\alpha t\}^2}.$$

For the same reason as in case (i), the infimum is taken over the set $x \neq \tilde{x}$ and $\alpha \in [-1, 1]$.

A.3 Proof of Theorem 4.4

We use the same notations as in the proofs of Theorems 4.1 and 4.3. The local critical radius $\theta_{c,loc}$ defined by (3.7) is rewritten as

$$\tan^2 \theta_{c,loc} = \liminf_{|x-\tilde{x}| \rightarrow 0, \alpha \rightarrow 1} \frac{(1 - \alpha s)^2}{1 - s^2 - \alpha^2 t^2}.$$

Let $\tilde{x} = x + \Delta$ and $\alpha = 1 - \delta$, and consider $\Delta \rightarrow 0$ and $\delta \rightarrow 0$. Write $\psi_x = \partial\psi/\partial x$, $\psi_{xx} = \partial^2\psi/\partial x^2$, etc., and $g = \psi_x^T \psi_x$, $\gamma = \psi_{xx}^T \psi_x$, and $\eta = \psi_{xx}^T \psi_{xx}$ as before. Noting that

$$\begin{aligned} 0 &= d(\psi^T \psi)/dx = 2\psi_x^T \psi, \\ 0 &= d^2(\psi^T \psi)/dx^2 = 2\psi_{xx}^T \psi + 2\psi_x^T \psi_x, \\ 0 &= d^3(\psi^T \psi)/dx^3 = 2\psi_{xxx}^T \psi + 6\psi_{xx}^T \psi_x, \\ 0 &= d^4(\psi^T \psi)/dx^4 = 2\psi_{xxxx}^T \psi + 8\psi_{xxx}^T \psi_x + 6\psi_{xx}^T \psi_{xx}, \end{aligned}$$

we have

$$\psi_{xx}^T \psi = -g, \quad \psi_{xxx}^T \psi = -3\gamma, \quad \psi_{xxxx}^T \psi + 4\psi_{xxx}^T \psi_x = -3\eta.$$

Substituting these, we have

$$\begin{aligned} s &= \psi(x)^T \psi(\tilde{x}) = \psi^T \left(\psi + \psi_x \Delta + \frac{1}{2} \psi_{xx} \Delta^2 + \frac{1}{6} \psi_{xxx} \Delta^3 + \frac{1}{24} \psi_{xxxx} \Delta^4 \right) + o(\Delta^4) \\ &= 1 - \frac{1}{2} g \Delta^2 - \frac{1}{2} \gamma \Delta^3 + \frac{1}{24} \psi_{xxxx}^T \psi \Delta^4 + o(\Delta^4), \\ r &= \psi_x(x)^T \psi(\tilde{x}) = \psi_x^T \left(\psi + \psi_x \Delta + \frac{1}{2} \psi_{xx} \Delta^2 + \frac{1}{6} \psi_{xxx} \Delta^3 \right) + o(\Delta^3) \\ &= g \Delta + \frac{1}{2} \gamma \Delta^2 + \frac{1}{6} \psi_{xxx}^T \psi_x \Delta^3 + o(\Delta^3), \\ t^2 &= \frac{r^2}{g} = g \Delta^2 + \gamma \Delta^3 + \frac{1}{4g} \gamma^2 \Delta^4 + \frac{1}{3} \psi_{xxx}^T \psi_x \Delta^4 + o(\Delta^4), \end{aligned}$$

and

$$\begin{aligned} 1 - s^2 - t^2 &= (1 - s)(2 - (1 - s)) - t^2 = 2(1 - s) - (1 - s)^2 - t^2 \\ &= 2 \left(\frac{1}{2} g \Delta^2 + \frac{1}{2} \gamma \Delta^3 - \frac{1}{24} \psi_{xxxx}^T \psi \Delta^4 \right) - \left(\frac{1}{2} g \Delta^2 \right)^2 \\ &\quad - \left(g \Delta^2 + \gamma \Delta^3 + \frac{1}{4g} \gamma^2 \Delta^4 + \frac{1}{3} \psi_{xxx}^T \psi_x \Delta^4 \right) + o(\Delta^4) \\ &= \frac{1}{4} \left(\eta - g^2 - \frac{\gamma^2}{g} \right) \Delta^4 + o(\Delta^4) = \frac{1}{4} \kappa g^2 \Delta^4 + o(\Delta^4), \end{aligned}$$

where

$$\kappa = \kappa(x) = \frac{\eta}{g^2} - \frac{\gamma^2}{g^3} - 1.$$

Note that κ is nonnegative because

$$0 \leq \det \begin{pmatrix} \psi^T \\ \psi_x^T \\ \psi_{xx}^T \end{pmatrix} \begin{pmatrix} \psi & \psi_x & \psi_{xx} \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & -g \\ 0 & g & \gamma \\ -g & \gamma & \eta \end{pmatrix} = \kappa g^3.$$

For the order of δ , we consider two cases: (i) $\delta/\Delta^2 \sim gc$ ($0 \leq c < \infty$) and (ii) $\Delta^2/\delta \sim 0$.

For case (i), noting that $\alpha = 1 - \delta$, $1 - s \sim \frac{1}{2} g \Delta^2$, $t^2 \sim g \Delta^2$,

$$\begin{aligned} (1 - \alpha s)^2 &= (1 - s + \delta - \delta(1 - s))^2 \sim (1 - s + \delta)^2 \sim \left(\frac{1}{2} g \Delta^2 + gc \Delta^2 \right)^2 = \frac{1}{4} g^2 (1 + 2c)^2 \Delta^4, \\ 1 - s^2 - \alpha^2 t^2 &= 1 - s^2 - t^2 + 2\delta t^2 - \delta^2 t^2 \sim \frac{1}{4} g^2 (\kappa + 8c) \Delta^4, \end{aligned}$$

and hence

$$\frac{(1 - \alpha s)^2}{1 - s^2 - k\alpha^2} \sim \frac{(1 + 2c)^2}{\kappa + 8c}. \quad (\text{A.10})$$

We consider the minimum value of (A.10) for $c \geq 0$. For $c > -\kappa/8$, (A.10) has a unique minimum value $1 - \kappa/4$ at $c = (2 - \kappa)/4$. Therefore,

$$\min_{c \geq 0} \frac{(1 + 2c)^2}{\kappa + 8c} = \begin{cases} 1 - \frac{\kappa}{4} & (\kappa \leq 2), \\ \frac{1}{\kappa} & (\kappa \geq 2). \end{cases}$$

For case (ii),

$$\begin{aligned} (1 - \alpha s)^2 &= (1 - s + \delta - \delta(1 - s))^2 \sim \delta^2, \\ 1 - s^2 - \alpha^2 t^2 &= 1 - s^2 - t^2 + 2\delta t^2 - \delta^2 t^2 \sim 2g\Delta^2\delta, \end{aligned}$$

and

$$\frac{(1 - \alpha s)^2}{1 - s^2 - k\alpha^2} \sim \frac{\delta}{2g\Delta^2} \rightarrow \infty.$$

In summary, we have

$$\tan^2 \theta_{c,loc} = \min \left\{ \inf_{x: \kappa(x) \leq 2} \left(1 - \frac{\kappa(x)}{4} \right), \inf_{x: \kappa(x) \geq 2} \frac{1}{\kappa(x)} \right\}.$$

A.4 Coverage probability under model misspecification

First, note that the best parameter under the model $\beta_i^T f(x)$ is given by $\beta_i^* = \Sigma X^T g_i$, where

$$X = \begin{pmatrix} f(x_1)^T \\ \vdots \\ f(x_n)^T \end{pmatrix}, \quad \Sigma = (X^T X)^{-1} = \left(\sum_{i=1}^n f(x_i) f(x_i)^T \right)^{-1}, \quad g_i = \begin{pmatrix} g_i(x_1) \\ \vdots \\ g_i(x_n) \end{pmatrix}.$$

Let

$$y_i = \begin{pmatrix} y_{i1} \\ \vdots \\ y_{in} \end{pmatrix}, \quad \varepsilon_i = \begin{pmatrix} \varepsilon_{i1} \\ \vdots \\ \varepsilon_{in} \end{pmatrix}.$$

The least square estimator for β_i^* is $\hat{\beta}_i^* = \Sigma X^T y_i$, which is distributed as $N_p(\beta_i^*, \Sigma)$. Let $b_{1-\alpha}$ be the threshold for $1 - \alpha$ bands when the assumed model is the true model. The approximate value of $b_{1-\alpha}$ can be obtained by the tube method.

On the other hand, when the true model is $g_i(x)$, the coverage probability becomes

$$\begin{aligned}
& P\left(\left|\sum_{i=1}^k c_i(\hat{\beta}_i^*)^T f(x) - \sum_{i=1}^k c_i g_i(x)\right| \leq b_{1-\alpha} \sqrt{f(x)^T \Sigma f(x)} \text{ for all } x \in \mathcal{X}, c \in \mathcal{C}\right) \\
&= P\left(\max_{x \in \mathcal{X}, c \in \mathcal{C}} \frac{\sum_{i=1}^k c_i(\hat{\beta}_i^*)^T f(x) - \sum_{i=1}^k c_i g_i(x)}{\sqrt{f(x)^T \Sigma f(x)}} \leq b_{1-\alpha}\right) \\
&= P\left(\max_{x \in \mathcal{X}, c \in \mathcal{C}} \left(\frac{\sum_{i=1}^k c_i(\hat{\beta}_i^* - \beta_i^*)^T f(x)}{\sqrt{f(x)^T \Sigma f(x)}} + \frac{\sum_{i=1}^k c_i((\beta_i^*)^T f(x) - g_i(x))}{\sqrt{f(x)^T \Sigma f(x)}}\right) \leq b_{1-\alpha}\right). \quad (\text{A.11})
\end{aligned}$$

Noting that, for the two functions $h_1(y)$ and $h_2(y)$ on \mathcal{Y} , if $\max_{y \in \mathcal{Y}}(-h_i(y)) = \max_{y \in \mathcal{Y}} h_i(y)$, then

$$\max_y h_1(y) \leq \max_y (h_1(y) + h_2(y)) + \max_y (-h_2(y)) = \max_y (h_1(y) + h_2(y)) + \max_y h_2(y),$$

hence,

$$\max_y h_1(y) - \max_y h_2(y) \leq \max_y (h_1(y) + h_2(y)) \leq \max_y h_1(y) + \max_y h_2(y).$$

Therefore, (A.11) is bounded below and above by $1 - P(b_{1-\alpha} - \delta)$ and $1 - P(b_{1-\alpha} + \delta)$, respectively, where

$$\begin{aligned}
\delta &= \max_{x \in \mathcal{X}, c \in \mathcal{C}} \frac{\sum_{i=1}^k c_i((\beta_i^*)^T f(x) - g_i(x))}{\sqrt{f(x)^T \Sigma f(x)}} \\
&= \max_{x \in \mathcal{X}} \sqrt{\frac{\sum_{i=1}^k ((\beta_i^*)^T f(x) - g_i(x) - \frac{1}{k} \sum_{i=1}^k ((\beta_i^*)^T f(x) - g_i(x)))^2}{f(x)^T \Sigma f(x)}}
\end{aligned}$$

as given in (5.2), and

$$P(b) = P\left(\max_{x \in \mathcal{X}, c \in \mathcal{C}} \frac{\sum_{i=1}^k c_i(\hat{\beta}_i^* - \beta_i^*)^T f(x)}{\sqrt{f(x)^T \Sigma f(x)}} \geq b\right).$$

An upper bound for the bias of coverage probability for a $1 - \alpha$ confidence band is

$$\max\{\alpha - P(b_{1-\alpha} + \delta), P(b_{1-\alpha} - \delta) - \alpha\},$$

which is approximated by

$$\Delta = \max\{\alpha - \hat{P}(\hat{b}_{1-\alpha} + \delta), \hat{P}(\hat{b}_{1-\alpha} - \delta) - \alpha\}$$

in (5.3), where $\hat{P}(b)$ is the tube approximation formula for $P(b)$ given in (4.1), and $\hat{b}_{1-\alpha}$ is the solution of $\hat{P}(b) = \alpha$.

Note that (A.11) is

$$P\left(\max_{x \in \mathcal{X}} \frac{\sum_{i=1}^k ((\hat{\beta}_i^*)^T f(x) - g_i(x) - \frac{1}{k} \sum_{i=1}^k ((\hat{\beta}_i^*)^T f(x) - g_i(x)))^2}{f(x)^T \Sigma f(x)} \leq b_{1-\alpha}^2\right),$$

which is used for the simulation study.

Acknowledgments

The authors thank Professors Toshihiko Shiroishi and Toyoyuki Takada of National Institute of Genetics, Japan, for providing the dataset of Takada, et al. (2008).

References

- Adler, R. J. and Taylor, J. E. (2007). *Random Fields and Geometry*, Springer.
- de Boor, C. (1978). *A Practical Guide to Splines*, Springer.
- Casella, G. and Hwang, J. T. G. (2012). Shrinkage confidence procedures, *Statistical Science*, **27** (1), 51–60.
- Faraway, J. J. and Sun, J. (1995). Simultaneous confidence bands for linear regression with heteroscedastic errors, *Journal of the American Statistical Association*, **90** (431), 1094–1098.
- Federer, H. (1959). Curvature measures, *Transactions of the American Mathematical Society*, **93**, 418–491.
- Hotelling, H. (1939). Tubes and spheres in n -spaces, and a class of statistical problems, *American Journal of Mathematics*, **61** (2), 440–460.
- Jamshidian, M., Liu, W., and Bretz, F. (2010). Simultaneous confidence bands for all contrasts of three or more simple linear regression models over an interval, *Computational Statistics and Data Analysis*, **54** (6), 1475–1483.
- Johansen, S. and Johnstone, I. M. (1990). Hotelling’s theorem on the volume of tubes: Some illustrations in simultaneous inference and data analysis, *The Annals of Statistics*, **18** (2), 652–684.
- Johnstone, I. and Siegmund, D. (1989). On Hotelling’s formula for the volume of tubes and Naiman’s inequality, *The Annals of Statistics*, **17** (1), 184–194.
- Kuriki, S. and Takemura, A. (2001). Tail probabilities of the maxima of multilinear forms and their applications, *The Annals of Statistics*, **29** (2), 328–371.
- Kuriki, S. and Takemura, A. (2009). Volume of tubes and the distribution of the maximum of a Gaussian random field, *Selected Papers on Probability and Statistics*, American Mathematical Society Translations Series 2, **227** (2), 25–48.
- Krivobokova, T., Kneib, T., and Claeskens, G. (2010). Simultaneous confidence bands for penalized spline estimators, *Journal of the American Statistical Association*, **105** (490), 852–863.

- Leeb, H., Pötscher, B. M., and Ewald, K. (2015). On various confidence intervals post-model-selection, *Statistical Science*, **30** (2), 216–227.
- Liu, W. (2010). *Simultaneous Inference in Regression*, Chapman & Hall/CRC.
- Liu, W., Jamshidian, M., and Zhang, Y. (2004). Multiple comparison of several linear regression models, *Journal of the American Statistical Association*, **99** (466), 395–403.
- Liu, W., Lin, S., and Piegorsch, W. W. (2008). Construction of exact simultaneous confidence bands for a simple linear regression model, *International Statistical Review*, **76** (1), 39–57.
- Liu, W., Wynn, H. P., and Hayter, A. J. (2008). Statistical inferences for linear regression models when the covariates have functional relationships: polynomial regression, *Journal of Statistical Computation and Simulation*, **78** (4), 315–324.
- Muirhead, R. J. (2005). *Aspects of Multivariate Statistical Theory*, 2nd ed., Wiley.
- Lu, X. and Chen, J. T. (2009). Exact simultaneous confidence segments for all contrast comparisons, *Journal of Statistical Planning and Inference*, **139** (8), 2816–2822.
- Naiman, D. Q. (1986). Conservative confidence bands in curvilinear regression, *The Annals of Statistics*, **14** (3), 896–906.
- Shen, X., Wolfe, D. A., and Zhou, S. (1998). Local asymptotics for regression splines and confidence regions, *The Annals of Statistics*, **26** (5), 1760–1782.
- Spurrier, J. D. (1999). Exact confidence bounds for all contrasts of three or more regression lines, *Journal of the American Statistical Association*, **94** (446), 483–488.
- Spurrier, J. D. (2002). Exact multiple comparisons of three or more regression lines: Pair-wise comparisons and comparisons with a control, *Biometrical Journal*, **44** (7), 801–812.
- Sun, J. (1993). Tail probabilities of the maxima of Gaussian random fields, *The Annals of Probability*, **21** (1), 34–71.
- Sun, J., and Loader, C. R. (1994). Simultaneous confidence bands for linear regression and smoothing, *The Annals of Statistics*, **22** (3), 1328–1346.
- Sun, J., Loader, C., and McCormick, W. P. (2000). Confidence bands in generalized linear models, *The Annals of Statistics*, **28** (2), 429–460.
- Sun, J., Raz, J., and Faraway, J. J. (1999). Confidence bands for growth and response curves, *Statistica Sinica*, **9**, 679–698.
- Takada, T., Mita, A., Maeno, A., Sakai, T., Shitara, H., Kikkawa, Y., Moriwaki, K., Yonekawa, H., and Shiroishi, T. (2008). Mouse inter-subspecific consomic strains for genetic dissection of quantitative complex traits, *Genome Research*, **18** (3), 500–508.
<http://genome.cshlp.org/content/18/3/500>

- Takemura, A. and Kuriki, S. (2002). On the equivalence of the tube and Euler characteristic methods for the distribution of the maximum of Gaussian fields over piecewise smooth domains, *The Annals of Applied Probability*, **12** (2), 768–796.
- Uusipaikka, E. (1983). Exact confidence bands for linear regression over intervals, *Journal of the American Statistical Association*, **78** (383), 638–644.
- Working, H. and Hotelling, H. (1929). Applications of the theory of error to the interpretation of trends, *Journal of the American Statistical Association*, **24** (165), 73–85.
- Weyl, H. (1939). On the volume of tubes, *American Journal of Mathematics*, **61** (2), 461–472.
- Worsley, K. (1995). Boundary corrections for the expected Euler characteristic of excursion sets of random fields, with an application to astrophysics, *Advances in Applied Probability*, **27** (4), 943–959.
- Wynn, H. P. and Bloomfield, P. (1971). Simultaneous confidence bands in regression analysis, *Journal of the Royal Statistical Society, Series B*, **33** (2), 202–217.